Structure and Function of the Homodimeric Reaction Center, and Hydrogen Production, in *Heliobacterium modesticaldum*

by

Christopher J. Gisriel

A Dissertation Presented in Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy

Approved October 2017 by the
Graduate Supervisory Committee:

Kevin Redding, Chair
Anne Jones
James Allen

ARIZONA STATE UNIVERSITY

December 2017

ABSTRACT

The evolution of photosynthesis caused the oxygen-rich atmosphere in which we thrive today. Although the reaction centers involved in oxygenic photosynthesis probably evolved from a protein like the reaction centers in modern anoxygenic photosynthesis, modern anoxygenic reaction centers are poorly understood. One such anaerobic reaction center is found in *Heliobacterium modesticaldum*. Here, the photosynthetic properties of *H. modesticaldum* are investigated, especially as they pertain to its unique photochemical reaction center.

The first part of this dissertation describes the optimization of the previously established protocol for the *H. modesticaldum* reaction center isolation. Subsequently, electron transfer is characterized by ultrafast spectroscopy; the primary electron acceptor, a chlorophyll *a* derivative, is reduced in ~25 ps, and forward electron transfer occurs directly to a 4Fe-4S cluster in ~650 ps without the requirement for a quinone intermediate. A 2.2-angstrom resolution X-ray crystal structure of the homodimeric heliobacterial reaction center is solved, which is the first ever homodimeric reaction center structure to be solved, and is discussed as it pertains to the structure-function relationship in energy and electron transfer. The structure has a transmembrane helix arrangement similar to that of Photosystem I, but differences in antenna and electron transfer cofactor positions explain variations in biophysical comparisons. The structure is then compared with other reaction centers to infer evolutionary hypotheses suggesting that the ancestor to all modern reaction centers could reduce mobile quinones, and that Photosystem I added lower energy cofactors to its electron transfer chain to avoid the formation of singlet oxygen.

In the second part of this dissertation, hydrogen production rates of *H. modesticaldum* are quantified in multiple conditions. Hydrogen production only occurs in cells grown without ammonia, and is further increased by removal of N_2. These results are used to propose a scheme that summarizes the hydrogen-production metabolism of *H. modesticaldum*, in which electrons from pyruvate oxidation are shuttled through an electron transport pathway including the reaction center, ultimately reducing nitrogenase. In conjunction, electron microscopy images of *H. modesticaldum* are shown, which confirm that extended membrane systems are not exhibited by heliobacteria.

DEDICATION

To my mother, whose humble and persistent example paved the way for all the success I have ever achieved.

To my family, whose patience knows no bounds.

To Savanah, with whom I walk this path.

To my Friends, who showed me a new way of living.

ACKNOWLEDGMENTS

Without the example of my Principal Investigator, Kevin Redding, exhibiting the level of enthusiasm for scientific research and dedication to it, I probably would have joined the endless procession of pre-meds. I whole-heartedly thank him, Patricia Baker, and the other members of the Redding Lab for making me feel like part of a family.

I am deeply grateful to those who selflessly guided me through some challenging periods in my graduate career, including Yuval Mazor, Shibom Basu, Chad Simmons, and many more. Your altruism is an inspiration and I will strive to extend to others the kindness you showed to me.

I thank the School of Molecular Sciences at Arizona State University, whose faculty are second to none, and who have provided me with a career that I love dearly. I am proud that my education has been molded by their profound expertise, especially by Anne Jones, Jim Allen, Neal Woodbury, Wade Van Horn, and the Frommes.

Lastly, I thank the Department of Energy, ASU's LightWorks, and ARCS Foundation for their financial contribution to my research.

TABLE OF CONTENTS

Page

CHAPTER Page

CHAPTER Page

APPENDIX Page

LIST OF TABLES

LIST OF FIGURES

Figure Page

Figure Page

Figure Page

LIST OF SCHEMES

Scheme Page

LIST OF ABBREVIATIONS

ATP	Adenosine triphosphate
BChl *g*	Bacteriochlorophyll *g*
CabRC	Chloroacidobacterial reaction center
CEF	Cyclic electron flow
CfxRC	Chloroflexi reaction center
Chl *a*	Chlorophyll *a*
CR	Charge recombination
CS	Charge separation
DADS	Decay-associated difference spectra
ET	Electron transfer
GC	Gas chromatography
GmRC	Gemmatimodetes reaction center
GsbRC	Green sulfur bacterial reaction center
HbRC	Heliobacterial reaction center
LGT	Lateral gene transfer

MQ	Menaquinone
MSA	Multiple sequence alignment
NAD	Nicotinamide adenine dinucleotide
ND	Non-decaying
NMOR	NADH:MQ oxidoreductase
PbRC	Purple bacterial reaction center
PEPCK	Phosphoenolpyruvate carboxykinase
PDB	Protein data bank
PFOR	Pyruvate:ferredoxin oxidoreductase
PGC	Photosynthetic gene cluster
PhQ	Phylloquinone
PSI	Photosystem I
PSII	Photosystem II
PYE	Pyruvate yeast extract
RC	Reaction center
ROS	Reactive oxygen species

TEM	Transmission electron microscopy
TCD	Thermal conductivity detector
TMH	Transmembrane helix
βDDM	*n*-dodecyl β-D-maltoside

Chapter 1

Introduction: An overview of photosynthesis and its diversity, biohydrogen production, and the synopsis and range of this dissertation

Christopher Gisriel[1]

[1]School of Molecular Sciences, Arizona State University, Tempe, AZ 85287

An Overview of Photosynthesis

Photosynthesis is the process by which the energy of light is converted to useful chemical energy by an organism to drive its metabolism. This is achieved by a variety of organisms in a variety of manners. In plants and cyanobacteria, oxygenic photosynthesis takes place in a membrane-rich organelle called the chloroplast. The chloroplast is composed of membrane stacks collectively called the thylakoid. This is where photosynthesis takes place. The thylakoid membranes are rich with transmembrane proteins that coordinate antenna for light harvesting. In oxygenic photosynthesis, two kinds of membrane proteins work in tandem to transport electrons and protons across the thylakoid membrane: the reaction centers (RCs) Photosystem I (PSI) and Photosystem II (PSII), the latter of which catalyzes the formation of oxygen from water. In contrast, anoxygenic organisms that perform photosynthesis use only one kind of RC. Other than PSI and PSII, the RCs are named simply after the organism in which they reside, like the purple bacterial reaction center (PbRC) and green sulfur bacterial reaction center (GsbRC). In anoxygenic photosynthesis, the RC catalyzes electron transfer (ET) that ultimately results in creating a proton-motive force across a membrane to drive adenosine triphosphate (ATP) synthesis [1]. In both oxygenic and anoxygenic photosynthesis, the first steps in this process involve light excitation and are appropriately named the "light reactions". In this set of reactions, photons are harvested by antenna molecules and this energy absorption leads to the transfer of electrons across a membrane, sometimes coupled with the pumping of protons for creating a charge gradient that can phosphorylate ADP to ATP and reduce NAD^+ or $NADP^+$. In the subsequent "dark

reactions" of oxygenic photosynthesis, these cofactors are used to drive carbon fixation, resulting in the production of carbohydrates, which are long-term energy storage molecules. The ultimate destination of the energy provided by the light reactions of photosynthesis, however, varies dependent upon the metabolic requirements of the organism.

Although many of the characteristics of photosynthesis between both organisms and individual RCs are preserved, billions of years of evolutionary divergence has resulted in a division of these traits among many species, metabolic pathways, and proteins. There is a wealth of information describing the functions of the RCs involved in oxygenic photosynthesis, but the photosynthesis performed by oxygen-intolerant organisms is significantly understudied and not well-understood, despite its importance to understanding how photosynthesis evolved in an anoxic world and adapted to oxygen when is became present in the atmosphere.

Reaction Center Function

RCs are membrane proteins that coordinate cofactors for energy transfer and ET. All RCs have a dimeric core which sandwich a chain of ET cofactors that span the membrane between the monomers. On the periphery of this dimer are antenna (bacteria)chlorophylls, which are the sites of photon absorption and energy transfer. Upon the absorption of a photon, energy is transferred between antenna pigments until it arrives at a special pair of pigments in the ET chain, often referred to as "P". P is the primary electron donor, the first reactive species in the ET chain. Upon excitation, P becomes strongly reducing. Other cofactors in the ET chain are coordinated in such a

manner that they are poised at redox potentials that form a gradient across the membrane, thereby transferring the electron through the chain from P, which is on one side of the membrane, to a terminal electron acceptor near the other side of the membrane. For clarity, this work will refer to the two sides of the membrane as the P-side (positive side, equivalent to the outer or donor side) and the N-side (negative side, equivalent to the cytoplasm or acceptor side). Therefore, the initial electron donor, "P", is closest to the P-side, and ET proceeds toward the N-side, where the terminal electron acceptors resides.

The terminal electron acceptor in the ET chain is commonly used to categorize RCs into two "types". Type I RCs employ a 4Fe-4S cluster (F_X) as their terminal electron acceptor, and Type II RCs employ a quinone as their terminal electron acceptor [2]. In Type I RCs, like PSI in oxidative photosynthesis and the heliobacterial reaction center (HbRC) in anoxygenic photosynthesis, a terminal 4Fe-4S cluster is oxidized by soluble ferredoxin driving downstream metabolic processes in the dark reactions. Whereas the ET mechanism of Type I RC's is a one-electron process, Type II RCs require a two-electron process. Type II RCs, like the PbRC and PSII in oxidative photosynthesis, employ a quinone as their terminal electron acceptor. Upon double reduction of the quinone and subsequent protonation, the fully reduced quinone, quinol, diffuses into the lipid bilayer to exchange with a fully oxidized quinone from the membrane's quinone pool, re-starting the process. Complete proton displacement across the membrane is achieved when the reduced quinone is re-oxidized at another transmembrane protein, the cytochrome *bc* complex, and the protons are released into the P-side of the membrane,

forming the proton gradient required for ATP synthesis by yet another membrane protein, ATP synthase.

Heliobacteria: Discovery and Classification

Heliobacteria were first discovered entirely by mistake. A microbiology class at Indiana University incorrectly prepared growth media for photoheterotrophic, nitrogen-fixing, purple non-sulfur bacteria, accidently replacing the NH_4Cl with NH_4SO_4 [3]. The sulfate was reduced to sulfide and inhibited the growth of the desired microorganism and greenish-brown mats were observed which were quite intolerant of oxygen, turning from brown to green in oxygen's presence, and dying shortly thereafter. Gest and Favinger pioneered the first experiments with the organism they then named *Heliobacterium chlorum* [4]. Heliobacteria were found to be unique; they contain bacteriochlorophyll *g* (BChl *g*), a pigment that is not found in any other organism, and they do not contain internal membranes or chlorosomes like those in photosynthetic green bacteria. 16S rRNA sequencing homology [5], the absence of lipopolysaccharide, and the high concentration of branched-chain fatty acids in the plasma membrane [6], led to its classification with Gram positive organisms, making it the first photosynthetic organism in its phylum, the Firmicutes. However, it has been suggested that heliobacteria ought to be classified with cyanobacteria because of their 16S rRNA sequence and RC homology to PSI [7].

Reaction Center Structures

Detailed understanding of RCs function requires knowledge of their structure. Structures of the 4 kinds of RCs and the years they were solved are shown in **Figure 1.1**. The first RC structure to be solved was that from the photosynthetic purple bacterium, *Rhodopseudomonas viridis*, via X-ray crystallography [8], for which a Nobel prize was awarded in 1988. This first RC structure has proved invaluable, serving as a stepping stone for the understanding of the structure – function relationship in membrane proteins in general, as well as photosynthetic RCs. Whereas the PbRC is considered the bacterial Type II RC, the structures of the RCs involved in oxidative photosynthesis, PSI and PSII, were solved in 1993 [9] and 2001 [10], respectively. Since these initial discoveries, the structures of these RCs have been refined to higher resolution and have been solved in a plethora of states from multiple species. These structures have been used to compare various features upon which a platform of study and discovery has arisen.

As mentioned above, all RC cores consist of a protein dimer that coordinates an ET chain. Each monomer of the dimer has 5 transmembrane helices (TMH) that surround the ET chain like a cage. Type II RCs are surrounded by peripheral membrane proteins that coordinate antenna pigments for excitation energy transfer to the core. The core protein of Type I RCs, however, has an additional 6 TMH that coordinate antenna molecules. Therefore, Type I RCs have 11 TMH: 6 N-terminal TMH coordinating antenna pigments, which we will refer to as the "antenna domain", and 5 C-terminal TMH coordinating the ET chain, which we will refer to as the "ET domain".

Although all RCs display similarities, including aspects of the ET chain, antenna molecules, and subunits, differences in RCs are a result of evolutionary divergence [11] and distinguish them from one another. Differences include, but are not limited to: the identities of (bacterio)chlorophylls utilized by the RC (the collection of which we will refer to as "(B)Chls" hereafter), the presence and arrangement of antenna-coordinating subunits and peripheral complexes, the nature of electron donors and acceptors to and from the RC, and whether the core of the RC is homodimeric or heterodimeric.

While the PbRC is considered the bacterial analog to PSII, because they are both Type II RCs, the HbRC and GsbRC are considered bacterial analogs to PSI. Of all known RCs, only the HbRC, GsbRC, and the RC from chloroacidobacteria (CabRC) are homodimeric. This trait is important when considering RC evolution, because a heterodimeric RC probably evolved from a homoodimeric RC [12,13]. All solved RC structures (including PSI, PSII and PbRC) are, however, heterodimeric, leaving a gap in our understanding of RC structures. No homodimeric RC structure has been solved until the work presented here, thus filling this gap, and providing insight into how photosynthesis may have evolved.

Evolution of Reaction Centers

Evolution is a difficult topic to study because of its complexity, the multiple scales on which to study it, and the fact that evolution occurs over billions of years during which intelligent life was unable to collect data, leaving us with only a single data point of evolution's state - now. The hope, however, lies in the likelihood that similarity in systems provides insight into the most likely characteristics of a common ancestor.

Although the ability to identify conservation in organisms, from their motility to their metabolism, gives insight into the evolution of photosynthesis, comparisons are complicated by the possibility of lateral gene transfers (LGT) and convergent evolution. Arguments that heliobacteria are especially primitive include their simple Gram-positive membrane architecture and their strict anaerobic requirement, making their habitat more similar to that of Earth's atmosphere before oxygen rose ~3 BYA [14]. Furthermore, it has been suggested that Heliobacteria are members of the most early-diverging photosynthetic lineage [15]. Those who study the evolution of photosynthesis on the nanoscale consider such topics as (bacterio)chlorophyll synthesis, antenna size and complexity, protein structure, and more. On the nanoscale, the HbRC is considered the most similar to the ancestor of all RCs for the following reasons: (1) it has a homodimeric core, (2) its core polypeptides share sequence homology to both Type I and Type II RCs, and (3) it has fewer subunits than any other RC. This argument is contradicted by an analysis of heliobacteria's pigment synthesis pathway[16,17]; heliobacteria are the only organism discovered thus far to make use of bacteriochlorophyll *g* (BChl *g*) [4], a pigment similar to chlorophyll *a* (Chl *a*) but exhibiting red-shifted absorption as a result of extra desaturation of the conjugated system in the porphyrin ring [18,19]. Chl *a* is an intermediate in the synthesis of BChl *g*, therefore the chlorophyll synthesis in heliobacteria requires an extra enzyme relative to organisms that make sole use of Chl *a*. This additional step argues that (bacterio)chlorophyll biosynthesis in heliobacteria is less primitive than in organisms that make use of only Chl *a*.

The evidence that the HbRC displays more "ancient" characteristics outweighs the evidence against it, but this is a constant subject of debate. All extant RCs diverged from a common ancestor [20] and as a result of the discoveries presented here, an additional layer of comparison can be gleaned concerning the evolutionary relationships between RCs, and predictions can be made regarding characteristics of the ancient ancestor of all RCs.

Heliobacterial Metabolism and Biohydrogen Production

Heliobacteria can grow photoheterotrophically in the light and chemotrophically in the dark [21]. Phototrophic growth requires the availability of organic carbon sources such as pyruvate or acetate and chemotrophic growth requires pyruvate for fermentation [22]. No CO_2-fixation pathways have been identified in heliobacteria previously [21]. Heliobacteria have been shown to fix nitrogen during both phototrophic and chemotrophic growth [21,22]. Like many Clostridia, their genome encodes for multiple hydrogenases and nitrogenases, both of which may produce hydrogen as a byproduct. Biohydrogen is hydrogen that is produced by algae, bacteria, and archea, and is sought as a potential alternative fuel. Molecular hydrogen from any source is a valuable resource, being used in industrial processes, such as oil-sands processing, coal gasification, and gas-line desulfurization. Hydrogen has also been considered a potential alternative fuel for replacing fossil fuels, but its use is limited by cost and other logistical factors. In Chapter 5, we quantify biohydrogen production by *H. modesticaldum*, and show that *H. modesticaldum* lack internal membranes or membrane invaginations by presenting

transmission electron microscopy (TEM) images of negatively-stained and thin-sectioned cells grown in various conditions.

Synopsis and Range of this Dissertation

The focus of this dissertation is to describe the characteristics of photosynthesis performed by *H. modesticaldum*, an anoxygenic phototrophic bacteria that uses the simplest RC known to perform the light reactions of photosynthesis, and compare the HbRC to the RCs of other organisms. Optimization of the purification procedure for the HbRC is described, and biophysical characterization of its activity is presented in Chapter 2. In Chapter 3, the first structure of a homodimeric RC is presented, an X-ray crystal structure of the HbRC at 2.2-Å resolution. A detailed structural analysis and comparison to other RCs is performed. Chapter 4 contains insight and hypotheses into how photosynthesis may have evolved, concluding that the ancestral RC could probably reduce mobile quinones, that the Type I/II split occurred as a result of inefficiency in quinol production, and that an ancestor of PSI evolved in response to oxygen exposure. Finally, the molecular hydrogen output from *H. modesticaldum* is quantified and optimized by manipulating the media in which it grows in Chapter 5. It concludes that heliobacteria are the first Gram-positive organism documented to perform photofermentation, and the first electron microscopy images of *H. modesticaldum* are shown, which provide insight into the organism's membrane architecture.

Figures

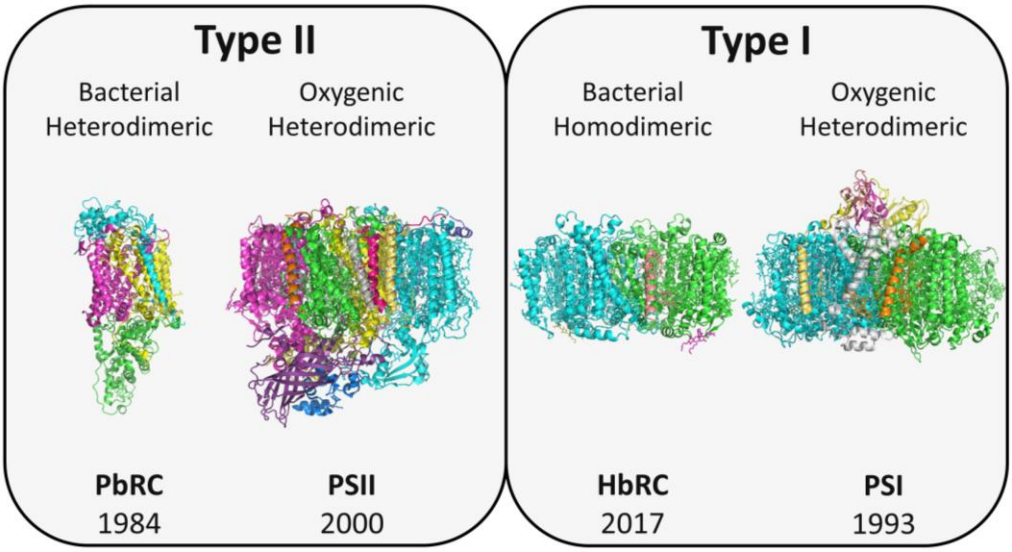


Figure 1.1 Structures of the 4 reaction centers. Type II (left) and Type I (right) RCs are categorized, and the year of the structure's publication is listed below their abbreviation. Each structure was imported from the PDB using PDB IDs 1PRC, 3WU2, 5V8K, and 1JB0 for the PbRC, PSII, HbRC, and PSI, respectively.

References

(1) Bryant, D. A.; Frigaard, N.-U. U. *Prokaryotic photosynthesis and phototrophy illuminated*; 2006; Vol. 14, pp 488–496.

(2) Schubert, W. D.; Klukas, O.; Saenger, W.; Witt, H. T.; Fromme, P.; Krauss, N. *J. Mol. Biol.* **1998**, *280* (2), 297.

(3) Gest, H. *Photosynth. Res.* **1994**, *41* (1), 17.

(4) Gest, H.; Favinger, J. L. *Arch. Microbiol.* **1983**, *136* (1), 11.

(5) Woese, C. R.; Debrunner-Vossbrinck, B. a; Oyaizu, H.; Stackebrandt, E.; Ludwig, W. *Science* **1985**, *229*, 762.

(6) Beck, H.; Hegeman, G. D.; White, D. *FEMS Microbiol. Lett.* **1990**, *69* (3), 229.

(7) Vermaas, W. F. J. *Photosynth. Res.* **1994**, *41* (1), 285.

(8) Deisenhofer, J.; Epp, O.; Miki, K.; Huber, R.; Michel, H. *Nature* **1985**, *318* (6047), 618.

(9) Krauss, N.; Hinrichs, W.; Witt, I.; Fromme, P.; Pritzkow, W.; Dauter, Z.; Betzel, C.; Wilson, K. S.; Witt, H. T.; Saenger, W. *Nature* **1993**, *361* (6410), 326.

(10) Zouni, A.; Witt, H. T.; Kern, J.; Fromme, P.; Krauss, N.; Saenger, W.; Orth, P. *Nature* **2001**, *409* (6821), 739.

(11) Blankenship, R. E. *Photosynth. Res.* **1992**, *33*, 91.

(12) Blankenship, R. E. *Molecular Mechanisms of Photosynthesis*, 2nd ed.; John Wiley & Sons, Ltd.: Southern Gate, Chicester, West Sussex, UK, 2008.

(13) Nitschke, W.; Rutherford, A. W. *Trends Biochem. Sci.* **1991**, *16* (7), 241.

(14) Fischer, W. W.; Hemp, J.; Johnson, J. E. *Annu. Rev. Earth Planet. Sci.* **2016**, *44* (1), 647.

(15) Gupta, R. S. *Photosynthesis Research*. 2003, pp 173–183.

(16) Bryant, D. A.; Liu, Z.; Li, T.; Zhao, F.; Costas, A. M. G.; Klatt, C. G.; Ward, D. M.; Frigaard, N.-U.; Overmann, J. In *Functional genomics and evolution of photosynthetic systems*; Springer Netherlands, 2012; pp 47–102.

(17) Sousa, F. L.; Shavit-Grievink, L.; Allen, J. F.; Martin, W. F. *Genome Biol. Evol.* **2013**, *5* (1), 200.

(18) Brockmann, H.; Lipinski, A.; Woese, C.; Madigan, M. *Arch. Microbiol.* **1983**, *136* (1), 17.

(19) Michalski, T. J.; Hunt, J. E.; Bowman, M. K.; Smith, U.; Bardeen, K.; Gest, H.; Norris, J. R.; Katz, J. J. *Proc. Natl. Acad. Sci. U. S. A.* **1987**, *84* (9), 2570.

(20) Cardona, T. *Photosynthesis Research*. 2015, pp 111–134.

(21) Tang, K.-H.; Yue, H.; Blankenship, R. E. *BMC Microbiol.* **2010**, *10*, 150.

(22) Kimble, L. K.; Madigan, M. T. *Arch. Microbiol.* **1992**, *158*, 155.

Chapter 2

Isolation and Characterization of the Heliobacterial Reaction Center

Christopher Gisriel[1], Dustin D. Luu[1], Su Lin[2], Kevin E. Redding[1]

[1]School of Molecular Sciences, Arizona State University, Tempe AZ 85287

[2]Biodesign Institute, Arizona State University, Tempe AZ 85287

Abstract

The heliobacterial reaction center is an ideal candidate for the study of reaction centers because of its simplicity, in that it lacks redox-active subunits other than the central dimeric core, and its electron transfer chain is composed of three unique (bacterio)chlorophyll species, unlike other reaction centers. Here, we modify its purification protocol to improve its oligomeric homogeneity confirmed by SDS-PAGE. Next, we verify that the reaction center purification maintained the activity known to be characteristic of its native state by using ultrafast transient absorption spectroscopy and ms-timescale photobleaching experiments. Ultrafast transient absorption spectroscopy identified the species involved in electron transfer, and the kinetics associated with redox events between them. An electrochromic bandshift of a bacteriochlorophyll *g* species was observed, implying that a bacteriochlorophyll *g* molecule lies between the primary electron donor and acceptor on each branch, similar to that observed in other RCs. Pump-probe photobleaching experiments on the ms timescale confirmed the charge recombination rate between the primary electron donor and terminal electron acceptor of ~15 ms as observed previously. The results of these experiments expanded upon and confirmed previously-observed characteristics of the heliobacterial reaction center, and enabled a purity to be achieved that was suitable for the crystallization experiments discussed in Chapter 3.

Introduction

The first heliobacterial reaction center (HbRC) purifications were performed by Trost and Blankenship, collecting the HbRC from sucrose density gradients to which

solubilized heliobacterial membranes had been applied and separated by ultracentrifugation [1]. The purification technique was later improved by replacing the sucrose density gradients with two ion exchange chromatography steps [2].

The reaction center (RC) core polypeptide was sequenced and found to be composed of a dimer of one polypeptide [3], now commonly termed PshA. Each PshA has a molecular weight of ~68 kDa [3]. The HbRC's main pigment is bacteriochlorophyll *g* (BChl *g*), which is unique to heliobacteria [4]. For ET, The HbRC is expected to contain two branches of ET cofactors that span the membrane by homology to other RCs. Because of its homodimeric oligomeric state, each branch is expected to appear identical to one another, displaying identical kinetics and spectral features during ET. The ET chain of the HbRC is unique in that it contains 3 distinct (bacterio)chlorophyll ((B)Chl) molecules. A generalized image of the (B)Chl species and the terms that designate their site names are shown in **Figure 2.1**. The primary electron donor, P_{800}, was previously determined to be composed of two BChl *g*' [5] and the primary electron acceptor, A_0, is thought to be an 8^1-OH Chl *a* [6]. Two more pigments of the same identity as those found in the bulk antenna of the RC, BChl *g*, are expected to lie between P_{800} and A_0 as seen in Photosystem I (PSI) [7]. These are often referred to as the "accessory" (B)Chls (Acc).

Forward ET proceeds to A_0 in ~25 ps and to F_X in ~650 ps [8,9]. Charge recombination (CR) between P_{800} and A_0 occurs in ~17 ns [10] and from F_X in ~15 ms [11]. Conflicting results for HbRC cofactors have been reported. Generally, however, the HbRC has been shown to contain the following cofactors: 16-30 BChl *g* [1,2,12], two BChl *g*' [5], two 8^1-OH Chl *a* [6], approximately one carotenoid (4,4' diaponeurosporene) [13], between one and two menaquinone-8 (MQ8) or menaquinone-9 (MQ9) [1,2], and one [4Fe-

4S] cluster [11]. (See our resolution of these conflicts in the X-ray crystallography experiments discussed in Chapter 3.)

Unlike all other known RCs, a quinone species is not required for ET [14]. Whether the HbRC employs a quinone in ET *in vivo*, however, has been a subject of continued debate [15–19]. Low levels of MQ can be achieved by extended washing of HbRC on the CM-Sepharose column during purification, although it appears to have no effect on the successful reduction of F_X following light-induced excitation of the RC. This is surprising because in PSI, the phylloquinone cofactor is absolutely required for forward ET to F_X [20]. The quinone content is estimated to be ~2.5 per RC in *H. mobilis* membranes [1] and ~4-5 per RC in *H. modesticaldum* membranes [2], both surprisingly low amounts. The quinone demand of other membrane proteins, like NADH dehydrogenase, fumarate reductase, and the cytochrome b_6c complex, coupled with the unaffected activity of the HbRC without quinones, leads us to hypothesize that quinones are not an essential part of the HbRC. Further, the crystal structure of the HbRC (see Chapter 3) does not contain a quinone. Quinones that have been reported may be nonspecifically-bound artifacts of purification, incorrectly assigned signals, or the quinone interaction could be of a transient nature, like that of the mobile quinone in Type II RCs.

The purification and characterization herein were performed for the purpose of optimizing purity toward facilitating successful crystal formation. However, the variations made to the purification protocol [2] (see Results and Discussion) should be considered as an optimized method for future HbRC purification.

Results and Discussion

Purification. The HbRC was isolated as published previously [2] with a modification to homogenize the resultant oligomeric state (see Materials and Methods). Previously, HbRC purification yielded heterogeneous oligomeric states (**Figure 2.2A**) but the addition of detergent following an ion-exchange chromatography step allowed for complete solubilization of the complex throughout the purification process. Whereas the HbRC was applied to the first column with ~0.9% *n*-dodecyl β-D-maltoside (βDDM), the stalled progression of the HbRC through the resin, likely as a result of non-specific interactions with the resin's substituents, caused a decrease in the detergent concentration as buffer exchange occurred with the equilibration buffer containing 0.02% βDDM. Therefore, after it was collected from the first column, HbRC was applied to the subsequent chromatography resin with a significantly lower detergent concentration than expected. This was the phenomenon that was suspected to have been causing heterogeneity of oligomeric states in the HbRC purification. The addition of detergent immediately after the first chromatography step allowed for maintained solubility, thus homogenous oligomeric state, of the HbRC population as it bound to the second column. This was confirmed by SDS-PAGE (**Figure 2.2B**). Further, the homogenous oligomeric state facilitated the growth of crystals used in X-ray crystallography experiments (see Chapter 3).

Characterization. The steady-state absorbance spectrum of the HbRC purification exhibits a typical combined (B)Chl spectrum (**Figure 2.3**). The major pigment, BChl *g*, has a Q_y peak at ~788 nm and a Q_x peak at ~570 nm. The small quantity of 8^1-OH Chl *a*

in the RC is apparent in a smaller Q_y peak around 670 nm. The Soret transitions of the various (B)Chls are seen in the UV/Blue region, below ~450 nm.

Ultrafast transient absorption experiments were used to map the progression of ET through the ET chain of the HbRC. These experiments were performed in the context of understanding the thermodynamics of ET to cofactors on the acceptor side of the heliobacterial membrane. In collaboration with the laboratory of John Golbeck (Penn State), multiple experiments were performed that identified the midpoint potential of the F_X/F_X^- couple to be ~-500 mV [9], which is ~175 mV less than the analogous [4Fe-4S] cluster in PSI. However, they also provided evidence consistent with the previously observed rate of forward ET to A_0 (decay time of ~25 ps) [8], suggesting that our preparation was fully active, despite the lack of quinones.

In both ascorbate- and dithionite-reduced HbRC samples, ultrafast transient absorption spectroscopy decay-associated difference spectra (DADS) showed an ~25-ps time component (black line in **Figures 2.4A and 2.4B**). This spectrum included a rise of bleaching centered at the Q_y peak of 8^1-OH Chl *a*, ~670 nm, and was assigned to the rise of A_0 photobleaching as is accepts an electron in initial charge separation (CS). What appears to be two peaks flanking the Q_y peak of BChl *g* at ~788 nm is a result of multiple processes. First, antenna BChl *g* that were excited prior to the first time point of the experiment are returning to their ground state. This creates a broad decay from ~740 to 850 nm. A second, sharper bleaching ~788 nm is due to oxidation of P_{800} in primary CS. Third, electrochromic bandshifts of the accessory and antenna BChl *g* near the two species becoming charged (P_{800} and A_0) add differential-like features. Without the

contribution from the decay of excited antenna, the 788 nm bleaching and electrochromic bandshift would alone appear as a positive peak at ~788 nm in the DADS. Because the DADS is the addition of multiple processes, it displays a missing section of the broad peak from antenna BChl *g* recovery, creating the shape of the DADS centered around the Q_y peak of BChl *g* in the ~25 ps component (**Figures 2.4A and 2.4B**).

The second decay component from the ascorbate-reduced HbRC sample, ~650 ps, shows a negative peak centered at the 8^1-OH Chl *a* Q_y peak, signifying the return of 8^1-OH Chl *a* to its ground (neutral) state. Because no other species was visibly identified, it was assumed that the electron was transferred to F_X. (The redox state of F_X cannot be determined, because the spectral region where it absorbs lies outside the range monitored in this experiment.) The ~650-ps decay component also contains a peak near the BChl *g* Q_y band, which is likely an electrochromic bandshift of the nearby accessory BChl *g* as the electron moves from A_0 to F_X. In the dithionite-reduced sample, the $P_{800}^+A_0^-$ difference spectrum was maintained as a non-decaying (ND) component that lasted longer than 3 ns, the length of the experiment. This is due to the pre-reduction of F_X by dithionite, rendering it unavailable for reduction by A_0. The ND component in the ascorbate-reduced sample shows only the bleached state of P_{800} as the $P_{800}^+F_X^-$ state persists for more than 3 ns.

Our ultrafast spectroscopy global analysis reveals a decay process where energy transfer in the antenna pool proceeds to the initial CS state (Ant/RC* $\rightarrow$ $P_{800}^+A_0^-$) in ~25 ps. Similarly, PSI's initial CS state is fully formed in ~24 ps [21]. In the PbRC, the initial CS state is achieved much faster, in ~3 ps, because of a lack of antenna pigments [22].

Upon purification, the HbRC is found to contain a low quinone content (see Appendix A for Chapter 3 Supplementary Information). It has been shown that upon quinone extraction, the HbRC maintains the same CR kinetics as before quinone extraction [14], making it quite different from PSI in that regard. It should be noted that the experiments presented here are unable to resolve quinone involvement as their characteristic absorption lies outside of the measured wavelength range. Another stark difference is that F_X is reduced in ~600 ps in the HbRC, whereas F_X in PSI is reduced in ~20 ns or ~200 ns, depending upon which branch's quinone is the electron donor [21]. Here, when the HbRC has F_X pre-reduced by dithionite, the ~650-ps decay component is lost, implying that F_X is unavailable to be reduced because of pre-reduction by dithionite.

The result that primary CS in the HbRC is so similar to that of PSI and that the reduction of F_X is so different than PSI must be due to a structural difference between the ET chains of the two RCs. This is unsurprising because the HbRC preparation used did not contain quinones (see Appendix A for Chapter 3 Supplementary Information), a cofactor required for ET to F_X in PSI. If no other cofactor is present between A_0 and F_X in the HbRC, PSI should exhibit a faster rate of ET from A_0 to the next acceptor in the ET chain. In PSI, the rate of ET from A_0 to PhQ is ~50 ps, much faster than forward ET from A_0 in the HbRC. The difference in rates suggest that the electron is transferred from A_0 a longer distance in the HbRC than in PSI, which is consistent with our hypothesis that ET occurs directly from A_0 to F_X. The longest possible distance for this rate can be calculated using the Moser-Dutton formula [23] (i.e. if □G = -□): $\log_{10}(k_{ideal}) = 15 - \square\square\square R$, where R is the closest distance between the cofactors. The 660-ps rate thus results in ~9.7 Å as the

longest possible distance between A_0 and F_X in the HbRC using this generalized formula. While this is an indirect method of analysis for calculating distances between cofactors, it can be deduced that the distance between A_0 and F_X in the HbRC must be shorter than the analogous cofactors in PSI. The distance between A_0 and F_X in the HbRC structure is directly measured in and compared to our ultrafast transient absorption data hypotheses in Chapter 3. Indeed, A_0 and F_X in the HbRC structure is shorter than the analogous distance in PSI by >2 Å.

Time-resolved photobleaching of P_{800} on the ms timescale displayed differences between CR of ascorbate- and dithionite-reduced HbRC. The P_{800} photobleaching seen in the dithionite-reduced sample had already dissipated before the first time point at 200 µs in the ms-timescale P_{800} photobleaching experiments, which is apparent by the low initial amplitude relative to the ascorbate-reduced sample. We attribute this to recombination from A_0 because of pre-reduced F_X before 200 µs. With F_X already reduced, CR occurs between P_{800}^+ and A_0^- in ~17 ns [10,24]. The ~15 ms photobleaching decay $t_{1/2}$ of P_{800} in the ascorbate-reduced sample (**Figure 2.5**) shows that the terminal electron acceptor, F_X, remains reduced for ~20 times the lifetime of the F_X cluster in PSI, which persists for about 0.5 - 1 ms [21]. Therefore, CR in the HbRC seems to be similar from A_0 but quite different from F_X, which can be explained by the higher midpoint potential of F_X in the HbRC, ~150 mV less reducing than F_X in PSI [9]. In PSI, the reduced F_X further reduces the F_A and F_B [4Fe-4S] clusters in PsaC, a peripheral subunit that extends into the acceptor-side soluble region to reduce soluble electron acceptors (*e.g.* ferredoxin, flavodoxin). In the HbRC, however, the long-lived stability of reduced F_X is beneficial

because it is the terminal electron acceptor in the ET chain and subsequent ET to soluble acceptors is dependent upon their diffusion, docking to the RC near F_X for direct reduction. In PSI, the terminal acceptor is F_A/F_B. Because these clusters perform the same function as F_X in the HbRC, they, too, exhibit a CR rate in the ms timescale, ~ 50-100 ms.

In summary, we have modified an isolation procedure of the HbRC that allows us to collect a pure product as a homogenous homodimeric RC. This homogeneity allowed for crystallization screens in which optimal conditions were identified. Characterization of the HbRC was performed to identify ET rates that were found to be consistent with previous measurements. In ultrafast transient absorption experiments, evidence for an accessory (B)Chl was detected, and the maximum distance of A_0 to F_X in the HbRC was hypothesized to be longer that the analogous distance in PSI.

Materials and Methods

Cell propagation and protein isolation. H. modesticaldum was grown anaerobically in pyruvate yeast extract (PYE) as previously described under 780 nm light at 51 °C for ~48 hours.[14] Cell growth and protein purification were performed as described previously to include membrane solubilization in 0.9% βDDM.[1]

To purify HbRC, solubilized membranes of *H. modesticaldum* were passed through a DEAE-Sepharose (anion exchange) column to remove PshB as described previously[1]. HbRC did not bind to the column and was collected as a dark brown elutant. Because previous purification attempts produced heterogeneous oligomeric states

(**Figure 2.2A**), we modified the protocol by increasing the detergent concentration before the next chromatography step. The βDDM concentration was assumed to be 0.2% (based on previous unpublished contact angle measurements of detergent concentrations by John Golbeck and Bryan Ferlez) and was increased to 2%. Then, the sample was applied to a CM Sepharose (cation exchange) column, where HbRC tightly bound to the top of the column. The HbRC was eluted with 100 mM $MgSO_4$[1]. A sample of the eluted product was analyzed via SDS-PAGE. A silver stain of the gel revealed a single band with an apparent molecular weight of ~47 kDa, which corresponds to the PshA subunit (**Figure 2.2**). In preparations made without including the extra solubilization step, minor bands at ~100 kDa and heavier were also present. In addition to producing homogenous product of dimeric PshA, an unexpected effect of modifying the βDDM concentration after anion exchange was that it allowed for successful crystallization (see below).

Pump-probe spectroscopy monitoring P_{800} photobleaching. For CR kinetic analysis, HbRC at a concentration of ~100 µM BChl *g* received a saturating excitation laser flash (6 ns, 15 mJ) at 523 nm (BioLogic JTS-10). Absorption was monitored at 810 nm (within the Q_y peak of BChl *g*) with 10 µs flashes. Upon excitation, instantaneous photobleaching was observed, implying a change in redox state of P_{800}. P_{800} photobleaching was fit with a monoexponential decay and $t_{1/2}$ were calculated.

Ultrafast transient absorption spectroscopy. Kinetic analysis of forward ET was performed via ultrafast transient absorption spectroscopy. A broad band pump-probe laser system described previously was used to perform ps timescale transient absorption measurements [25]. A regenerative amplifier system generated 100 fs laser pulses at 800 nm

(1 kH, Tsunami and Spitfire, Spectra-Physics), part of which was used to form 575 nm excitation pulses. The broad-band probe was formed by directing an 800 nm beam into a 3-mm sapphire plate which allowed the beam to be focused into an optical compressor before completing its path to the sample. White light probes were dispersed by a spectrophotometer using a CCD camera (DU420, Andor Technology). Resolution of thr data was at ~2.3 nm. Pump polarization was set to 54.7° to probe pulses. Two samples of HbRC were examined in this manner. The first contained Na-ascorbate (50 mM) to reduce residual P_{800}^{+} in the sample prior to excitation and the second contained Na-dithionite (50 mM) which also pre-reduced F_X. The ascorbate-reduced sample was contained in a 1.2-mm spinning wheel cuvette (OD_{788} = ~1.5) and the dithionite-reduced sample was contained in a 2-mm anaerobic quartz cuvette. Global fitting to a multi-exponential kinetic model were used to create decay-associated difference spectra (DADS), which were calculated and fit using ASUFIT (http://www.public.asu.edu/~laserweb/asufit/asufit.html) and resultant components were assigned to various ET events. For the sample containing ascorbate (**Table 2.1**, **Figure 2.4**), the DADS yielded two decay components and a non-decaying (ND) component.

Acknowledgements

This work was funded by the Division of Chemical Sciences, Geosciences, and Biosciences, Office of Basic Energy Sciences, of the U.S. Department of Energy through Grant (DE-SC0010575 to KR, and JHG).

Figures

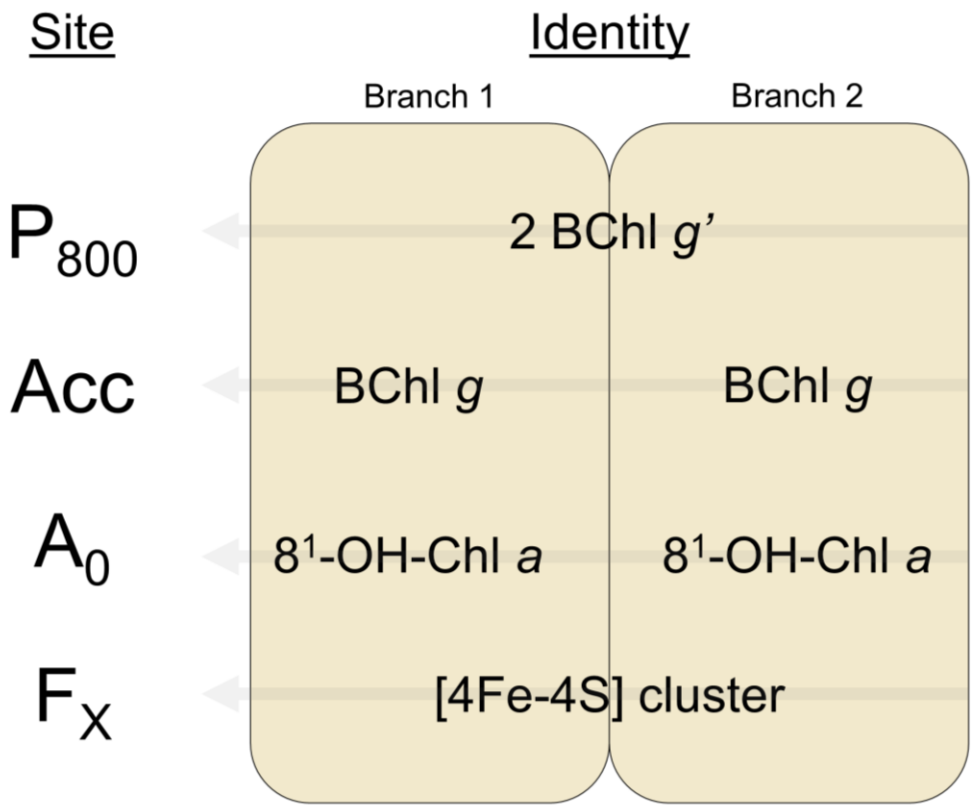


Figure 2.1 Overall arrangement of ET cofactors in the HbRC. Site names are listed on the left and the identities of the (B)Chl species and the terminal metal cluster are shown in a generalized orientation on the right.

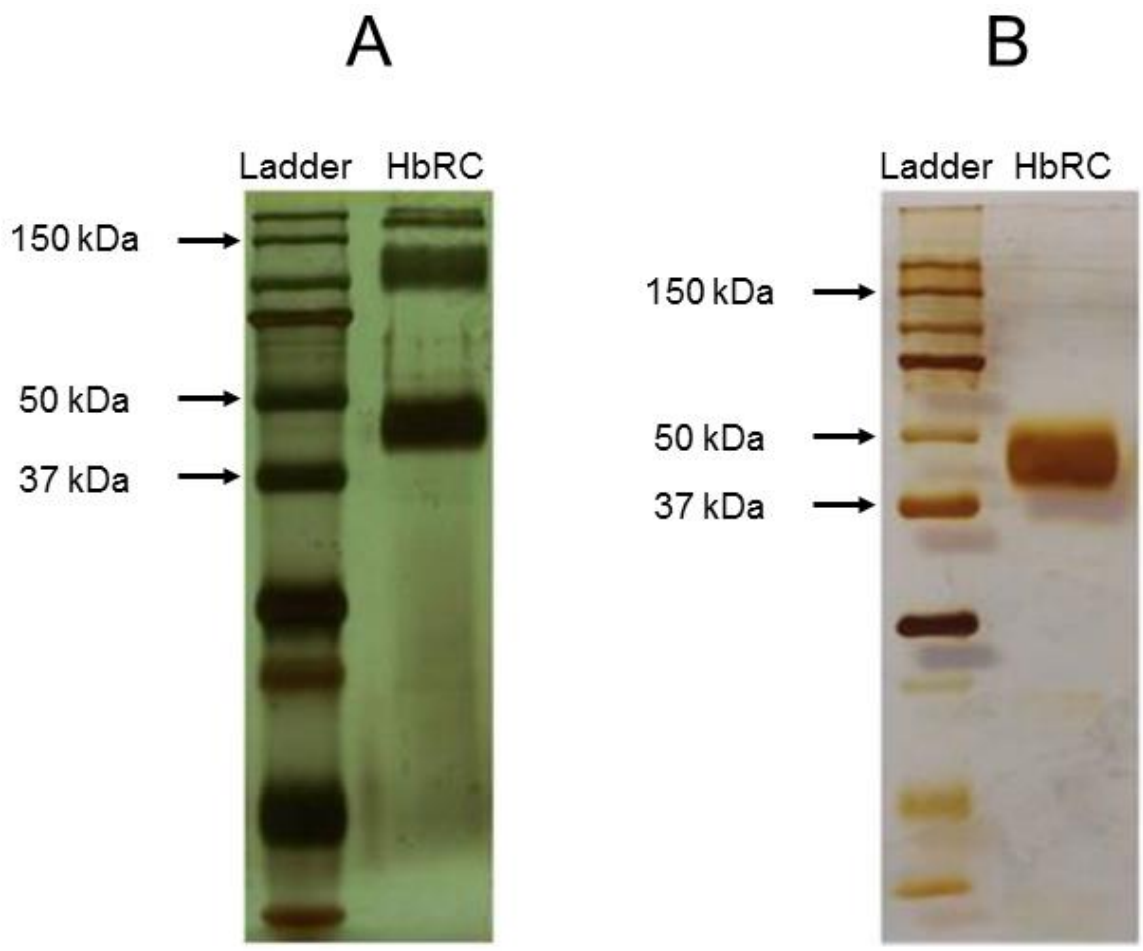


Figure 2.2 Silver-stained SDS-PAGE of purified HbRC. **A** shows SDS-PAGE of the HbRC purification before the extra detergent step was added between chromatography steps (right lane), relative to a standard ladder (left lane). **B** shows SDS-PAGE of the HbRC purification after the extra detergent step was added between chromatography steps (right lane), relative to a standard ladder (left lane). Both gels were made with 12% acrylamide and loaded with protein normalizing to ~3 µg BChl *g*. Labels indicate relevant molecular weight markers.

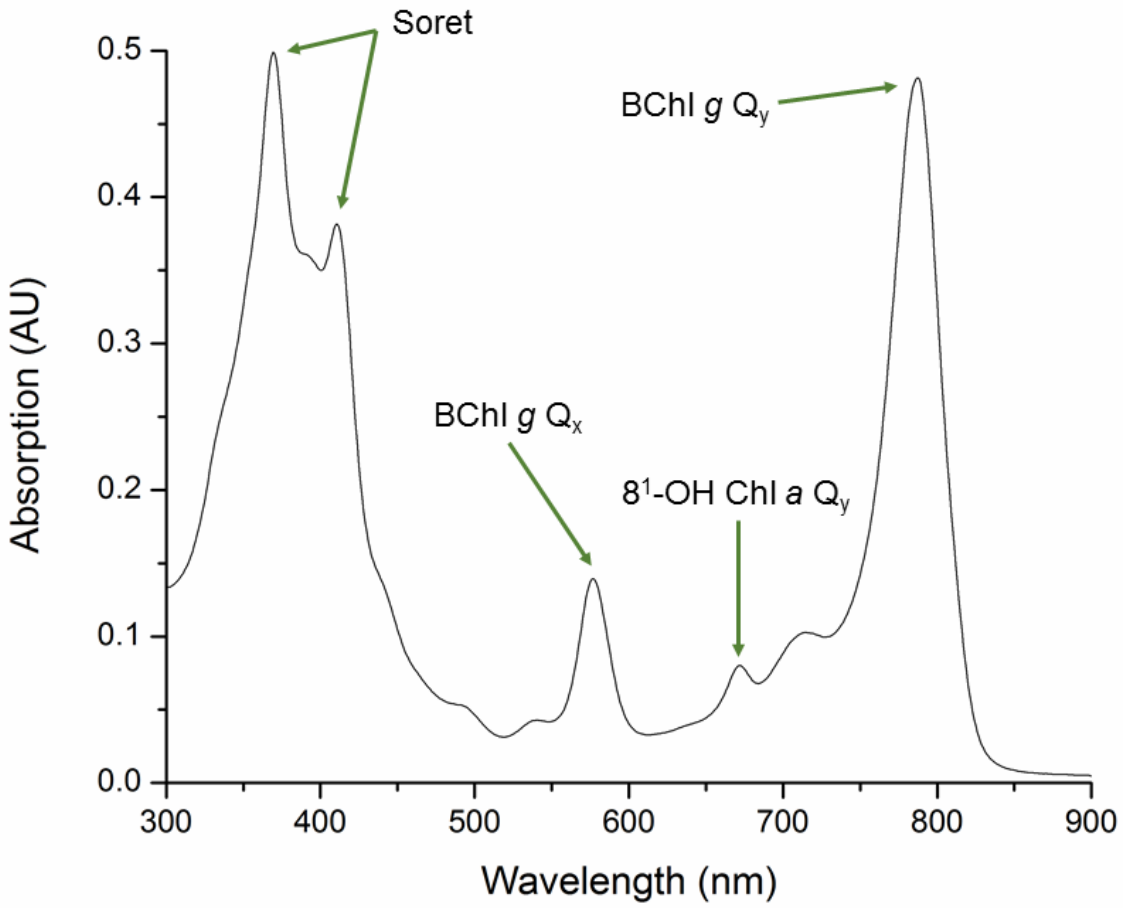
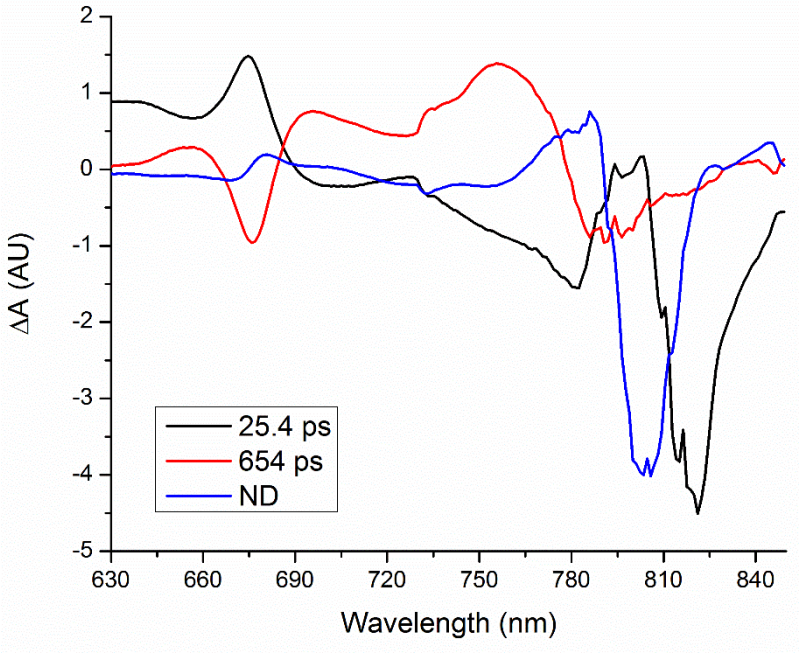


Figure 2.3 Absorbance spectrum of an HbRC isolation. Apparent peaks from BChl *g* include those at 368, 409, 575, and 788 which correspond to the two Soret transitions, Q_x band and Q_y band respectively. The 8^1-OH Chl *a* contributes a small peak in its Q_y band at 670 nm.

A



B

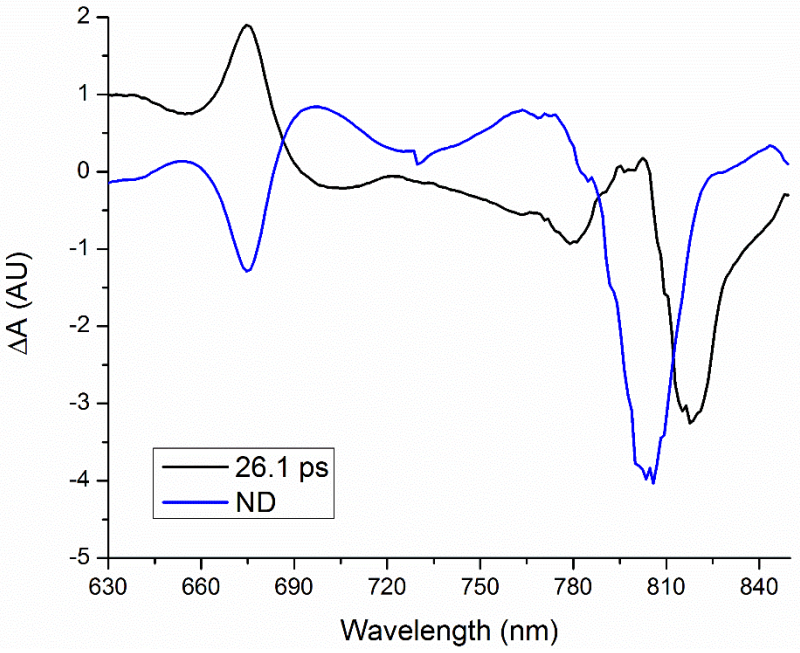


Figure 2.4 Ultrafast transient absorption of HbRC. DADS of HbRC

containing ascorbate is shown in panel **A** and dithionite in panel **B**. Data from HbRC containing ascorbate were fit with two exponential decay components with decay times of ~25 ps (blue) and ~654 ps (red) with a non-decaying (ND) component (gray). Data from HbRC containing dithionite was fit with a single exponential decay component of ~26 ps (blue) and a ND component (black).

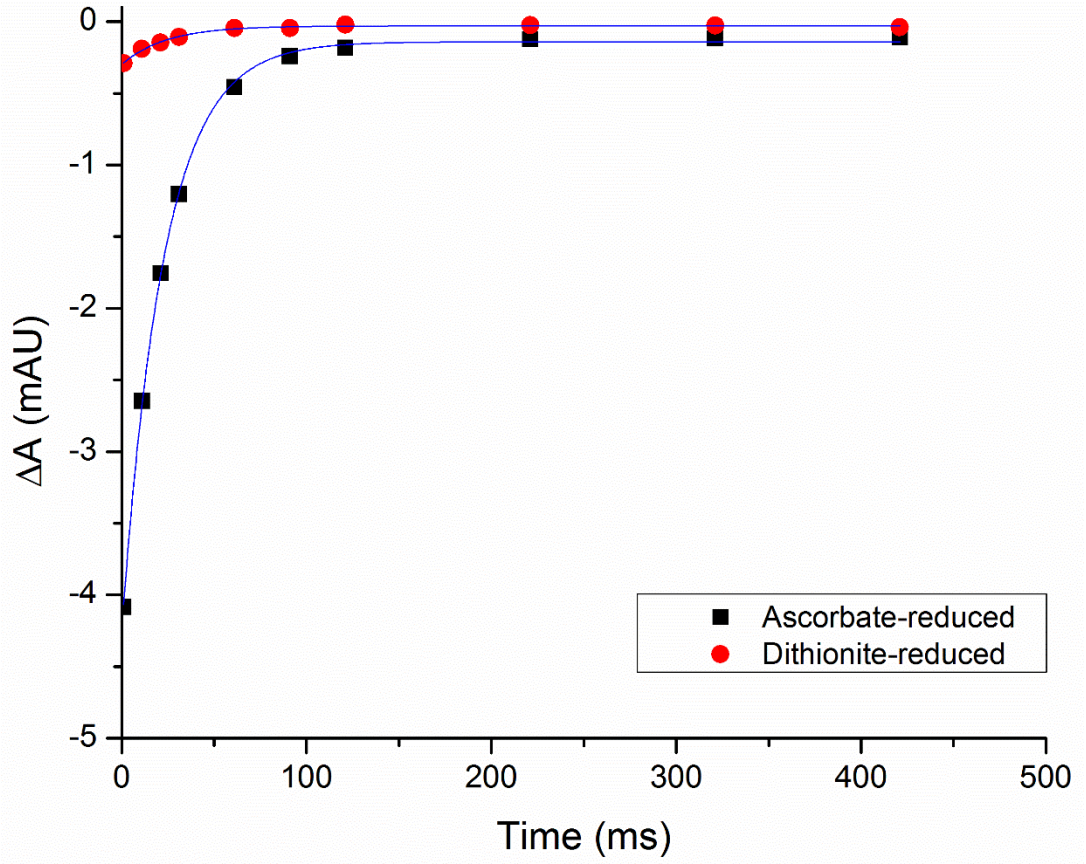


Figure 2.5 Millisecond-timescale P_{800} photobleaching. Data sets are overlaid and contain HbRC at a concentration of ~100 µM BChl *g*. Black squares represent the data points from the ascorbate-reduced HbRC and red circles represent the data points from the dithionite-reduced HbRC. Monoexponential decay fits to each data set are shown as blue lines. The $t_{1/2}$ of both decays was ~15 ms.

Tables

Table 2.1 Assignments of time components from DAS in HbRC with Na-ascorbate or Na-dithionite

Ascorbate

Decay Time	Assignment
25	Ant/RC* $\rightarrow P_{800}^{+}A_0^{-}$
654	$P_{800}^{+}A_0^{-} \rightarrow P_{800}^{+}F_X^{-}$
ND	Long-lived $P_{800}^{+}F_X^{-}$

Dithionite

Decay Time	Assignment
25	Ant/RC* $\rightarrow P_{800}^{+}A_0^{-}$
ND	Long-lived $P_{800}^{+}A_0^{-}$

References

(1) Trost, J. T.; Blankenship, R. E. *Biochemistry* **1989**, *28* (26), 9898.

(2) Sarrou, I.; Khan, Z.; Cowgill, J.; Lin, S.; Brune, D.; Romberger, S.; Golbeck, J. H.; Redding, K. E. *Photosynth. Res.* **2012**, *111* (3), 291.

(3) Liebl, U.; Mockensturm-Wilson, M.; Trost, J. T.; Brune, D. C.; Blankenship, R. E.; Vermaas, W. F. J. *Proc. Natl. Acad. Sci. U. S. A.* **1993**, *90* (15), 7124.

(4) Brockmann, H.; Lipinski, A.; Woese, C.; Madigan, M. *Arch. Microbiol.* **1983**, *136* (1), 17.

(5) Kobayashi, M.; Watanabe, T.; Ikegami, I.; van de Meent, E. J.; Amesz, J. *FEBS Lett.* **1991**, *284* (1), 129.

(6) van de Meent, E. J.; Kobayashi, M.; Erkelens, C.; van Veelen, P. A.; Amesz, J.; Watanabe, T. *BBA - Bioenerg.* **1991**, *1058* (3), 356.

(7) Krauss, N.; Hinrichs, W.; Witt, I.; Fromme, P.; Pritzkow, W.; Dauter, Z.; Betzel, C.; Wilson, K. S.; Witt, H. T.; Saenger, W. *Nature* **1993**, *361* (6410), 326.

(8) Chauvet, A.; Sarrou, J.; Lin, S.; Romberger, S. P.; Golbeck, J. H.; Savikhin, S.; Redding, K. E. *Photosynth. Res.* **2013**, *116* (1), 1.

(9) Ferlez, B.; Cowgill, J.; Dong, W.; Gisriel, C.; Lin, S. S.; Flores, M.; Walters, K.; Cetnar, D.; Redding, K. E. K. E.; Golbeck, J. H. J. H. *Biochemistry* **2016**, *55* (16), 2358.

(10) Kleinherenbrink, F. A. M.; Aartsma, T. J.; Amesz, J. *BBA - Bioenerg.* **1991**, *1057* (3), 346.

(11) Heinnickel, M.; Agalarov, R.; Svensen, N.; Krebs, C.; Golbeck, J. H.; Park, U. V; Pennsyl, V. *Biochemistry* **2006**, *45* (21), 6756.

(12) Nuijs, A. M.; van Dorssen, R. J.; Duysens, L. N. M.; Amesz, J. *Proc. Natl. Acad. Sci. U. S. A.* **1985**, *82* (20), 6865.

(13) Takaichi, S.; Inoue, K.; Akaike, M.; Kobayashi, M.; Oh-oka, H.; Madigan, M. T. *Arch. Microbiol.* **1997**, *168* (4), 277.

(14) Kleinherenbrink, F. A. M.; Ikegami, I.; Hiraishi, A.; Otte, S. C. M.; Amesz, J. *Biochim. Biophys. Acta - Bioenerg.* **1993**, *1142* (1–2), 69.

(15) Brok, M.; Vasmel, H.; Horikx, J. T. G.; Hoff, A. J. *FEBS Lett.* **1986**, *194* (2), 322.

(16) Muhiuddin, I. P.; Rigby, S. E. J.; Evans, M. C. W.; Amesz, J.; Heathcote, P. *Biochemistry* **1999**, *38* (22), 7159.

(17) Brettel, K.; Leibl, W.; Liebl, U. *Biochimica et Biophysica Acta - Bioenergetics*. 1998, pp 175–181.

(18) Miyamoto, R.; Mino, H.; Kondo, T.; Itoh, S.; Oh-oka, H. *Biochemistry* **2008**, *47* (15), 4386.

(19) Kondo, T.; Matsuoka, M.; Azai, C.; Mino, H.; Oh-oka, H.; Itoh, S. Springer, Berlin, Heidelberg, 2013; pp 21–24.

(20) McConnell, M. D.; Sun, J.; Siavashi, R.; Webber, A.; Redding, K. E.; Golbeck, J. H.; Van Der Est, A. *Biochim. Biophys. Acta - Bioenerg.* **2015**, *1847* (4–5), 429.

(21) Brettel, K.; Leibl, W. *Biochimica et Biophysica Acta - Bioenergetics*. 2001, pp 100–114.

(22) Lancaster, C. R. D.; Michel, H.; Lancaster, C. R. D.; Michel, H. In *Handbook of Metalloproteins*; John Wiley & Sons, Ltd: Chichester, 2006.

(23) Page, C. C.; Moser, C. C.; Chen, X.; Dutton, P. L. *Nature* **1999**, *402*, 47.

(24) Kleinherenbrink, F. A. M.; Amesz, J. *BBA - Bioenerg.* **1993**, *1143* (1), 77.

(25) Pan, J.; Lin, S.; Allen, J. P.; Williams, J. C.; Frank, H. A.; Woodbury, N. W. *J. Phys. Chem. B* **2011**, *115* (21), 7058.

Chapter 3

Structure of a Symmetric Photosynthetic Reaction Center-Photosystem

Christopher Gisriel[1], Iosifina Sarrou[2], Bryan Ferlez[3], John H. Golbeck[3,4], Kevin E. Redding[1,5], Raimund Fromme[1,6*]

[1]School of Molecular Sciences, Arizona State University, Tempe AZ 85287

[2]Center for Free-Electron Laser Science, DESY, 22607 Hamburg, Germany

[3]Department of Biochemistry and Molecular Biology, The Pennsylvania State University, University Park, Pennsylvania 16802

[4]Department of Chemistry, The Pennsylvania State University, University Park, Pennsylvania 16802

[5]Center for Bioenergy & Photosynthesis, Arizona State University

[6]Center of Applied Structural Discovery, Biodesign Institute, Tempe AZ 85287

Reproduced with permission from:

Gisriel, C.; Sarrou, I.; Ferlez, B.; Golbeck, J. H.; Redding, K. E.; Fromme, R. *Science.* **2017**, *357* (6355), 1021.

Abstract

Reaction centers are pigment-protein complexes that drive photosynthesis by converting light into chemical energy. It is believed that they arose once from a homodimeric protein. The symmetry of a homodimer is broken in heterodimeric reaction center structures, such as those reported previously. The 2.2-Å resolution X-ray structure of the homodimeric reaction center-photosystem from the phototroph *Heliobacterium modesticaldum* exhibits perfect C_2 symmetry. The core polypeptide dimer and two small subunits coordinate 54 bacteriochlorophylls and 2 carotenoids that capture and transfer energy to the electron transfer chain at the center, which performs charge separation and consists of 6 (bacterio)chlorophylls and an iron-sulfur cluster; unlike other reaction centers, it lacks a bound quinone. This structure preserves characteristics of the ancestral reaction center, providing insights into evolution of photosynthesis.

Background and Significance

Photosynthesis has changed the atmosphere of our planet and is the single most important energy conversion process driving the biosphere [1]. The light reactions of photosynthesis involve the conversion of light energy to chemical energy and are catalyzed by specialized membrane proteins called reaction centers (RC). Photons are absorbed by antenna molecules and the excitation energy is transferred to the core of the reaction center, where the excited state undergoes a charge separation process in which an electron is transferred from one bacteriochlorophyll (BChl), chlorophyll (Chl), or pheophytin to another. The electron is then transferred through a chain of cofactors to

drive downstream metabolism. RCs are typically classified by their terminal electron acceptor: either a [4Fe-4S] cluster (Type I) or a quinone (Type II) [2].

Oxygenic photosynthesis in higher plants, green algae and cyanobacteria makes use of Photosystem I (PSI), which is a Type I RC, and Photosystem II (PSII), which is a Type II RC. These work together to transfer electrons from water to ferredoxin. By contrast, anoxygenic phototrophic bacteria, such as *Heliobacterium modesticaldum*, use a single RC to drive a cyclic electron transfer (ET) pathway that creates a proton-motive force across the membrane, which is used to drive adenosine 5'-triphosphate (ATP) synthesis [3]. The discovery of heliobacteria [4] led to the identification of unique characteristics of its RC [5,6], now visualized in the structure herein. The heliobacterial RC has its antenna domain fused to the ET domain like PSI. For historic reasons, we maintain the term RC for the entire complex and distinguish its two domains: the ET domain and the antenna domain.

Despite the low sequence homology of core RC polypeptides from different organisms, reflecting billions of years of evolutionary divergence, all of these proteins share structural similarities within their central ET domain [7]. This region is at the interface of the two core subunits, each of which contributes five transmembrane helices (TMHs) that enclose the ET cofactors like a cage. The Type II RC from anoxygenic purple proteobacteria (PbRC) is heterodimeric and relies on peripheral antenna subunits to harvest photons. The core polypeptides of PSI (PsaA and PsaB) have an additional six TMHs at the N-terminal side that form a domain coordinating the bulk of the antenna pigments. The core polypeptides of anoxygenic phototrophs that use a Type I RC –

heliobacteria, chlorobia, and chloroacidobacteria – have a similar arrangement: an N-terminal antenna domain (six TMHs) and C-terminal ET domain (five TMHs). However, as the genomes of these organisms encode a single core RC polypeptide, their RCs should be homodimeric [8–10]. A homodimeric RC almost certainly preceded heterodimeric RCs in evolution [11]. Duplication of the core RC-subunit gene followed by divergence of the two genes would allow for conversion of a homodimeric RC to a heterodimeric RC. This likely occurred on at least 3 separate occasions, leading to the creation of PSI, PSII, and the RCs of purple bacteria and Chloroflexi [7]. High-resolution structures have been obtained from multiple heterodimeric RCs [PbRC, PSI, and PSII [12–14]], but no homodimeric RC structures have been solved until now.

Here, we present the first structure of a homodimeric Type I RC, an x-ray crystal structure at 2.2 Å resolution (crystallography statistics are presented in **Table A1**). It was purified from *H. modesticaldum*, a thermophilic anaerobe isolated from volcanic soil in Iceland [15]. Heliobacteria are the only phototrophic family in the Firmicutes (i.e. low-guanine-cytosine (GC) Gram-positive bacteria) and the only organisms that use bacteriochlorophyll *g* (BChl *g*), an isomer of chlorophyll *a* (Chl *a*). Despite the low-light habitat of heliobacteria, peripheral antenna complexes have yet to be identified. The RC used by heliobacteria is the only RC that does not seem to require a tightly bound quinone in ET [16], a characteristic that is supported by the structure presented herein. Most commonly referred to as the heliobacterial RC (HbRC), this complex is the simplest known RC, and it has been proposed to be the closest homolog to the common ancestor of all photosynthetic RCs [17,18].

Overall structure of the HbRC

The HbRC is elliptical in cross section, with major and minor axes of ~114 and 58 Å, respectively (**Figure 3.1B**). The average thickness of the hydrophobic membrane core is predicted to be ~27 Å [19] (see supplementary materials). The average height of the HbRC is ~50 Å (**Figure 3.1A**). The complex contains 24 TMHs (**Figure 3.2A**): 22 TMHs from the PshA homodimer and two more from the newly identified PshX subunits (see supplementary materials and **Figure A1** for details of its discovery). The HbRC binds 54 BChl *g*, 4 BChl *g*', two 8^1-hydroxychlorophyll *a* (8^1-OH Chl a_F), 2 carotenoids (4,4'-diaponeurosporene), 2 lipids, and 1 [4Fe-4S] cluster; four of these cofactors had not previously been described in the Protein Data Bank (see **Figure A2**). The PshA and PshX polypeptides account for ~140 kDa of the homodimeric RC. The total molecular mass is ~200 kDa, owing to the high cofactor content.

As a consequence of the homodimeric core, the ET cofactors are arranged into two identical branches about the C_2 symmetry axis (**Figure 3.3**). The first ET cofactor (P_{800}) is a pair of BChl *g*' molecules, the stereoisomer at the 13^2 position of ring E of BChl *g* [20], and is located toward the positive side (P-side, equivalent to the outside) of the membrane. On the negative side (N-side, equivalent to the cytoplasm), the segment between TMH8 and 9 of each PshA has two conserved Cys residues that ligate F_X, the terminal [4Fe-4S] cluster. Between P_{800} and F_X on each symmetric branch are a BChl *g* and an 8^1-OH Chl a_F.

Antenna. The antenna BChls forms two layers within the RC (dotted gray boxes in **Figure A3A**), one closer to the P-side and one closer to the N-side (grey boxes in

Figure A2A). Most are relatively close to one another (*i.e.*, within 6 Å), which would allow rapid excitation-energy-transfer among them. The bulk of this population (blue in **Figure 3.2B**) is ≥14 Å away from the ET chain (red in **Figure 3.2B**). There are only three antenna pigments on each side that are within 13 Å of any BChls and Chl in the ET chain (teal in **Figure 3.2B**). Energy transferred to the ET chain from the antenna pool probably arrives via one of these six BChl *g* molecules.

Most of the antenna BChls (50 of 54) are coordinated by the two PshA polypeptides, primarily from the six TMHs of the antenna domain (**Figures 3.1 and 3.2**); the other four BChl *g* are coordinated by the PshX subunits. Of the 27 BChls and 1 Chl associated with each PshA polypeptide, 19 are coordinated by His sidechains, 1 by Asn, 1 by Glu, 1 by Gln, 5 by water molecules, and 1 by an unidentified atom that has higher electron density than a water molecule (purple sphere in **Figures 3.3 and 3.4B**). Both BChls associated with PshX are both His-coordinated. Surprisingly, there is a BChl *g*' associated with each PshA within the center of an antenna pigment cluster.

ET Chain. A notable feature of the ET chain is that it contains three chemically distinct BChls or Chls. It begins on the P-side of the HbRC with P_{800}, a BChl *g*' dimer that serves as the primary electron donor, and it ends on the N-side with F_X, a [4Fe-4S] cluster serving as the terminal acceptor (**Figure 3.3**). Each BChl *g*' in P_{800} is coordinated by His537 from TMH10, whereas F_X is ligated by Cys432 and Cys441 from the loop between TMH8 and 9 of each PshA. The primary donor (P_{800}) and terminal acceptor (F_X) are located on the C_2 symmetry axis. There are two pairs of BChls or Chls on each side of the symmetry axis that link P_{800} to F_X. The one closer to P_{800} is a BChl *g*, while the one

closer to F_X is an 8^1-OH Chl a_F, in agreement with modeling of the electrochromic bandshift of the 8^1-OH Chl a_F spectrum in the presence of oxidized P_{800}^+ [21].

On the basis of biochemical and biophysical measurements [21–23], it was expected that the primary acceptor in the HbRC, called A_0 by analogy with the corresponding cofactor in PSI, was an 8^1-OH-Chl a_F molecule. The cofactor between the special pair and primary acceptor is often referred to as the "accessory" BChl or Chl (Acc). We will use that term here, but it should not be taken as a functional designation, as the analogous cofactor has been suggested to serve as the primary electron donor in PSI [24,25], PSII [26], and the PbRC [27–29]. Acc is coordinated by a small molecule approximately the size of water, and A_0 is coordinated by a water molecule. These are H bonded by the sidechains of Gln458 and Ser545, respectively (see discussion later in text). As in the other RCs, within a branch of the ET chain, the special pair (P_{800}) and primary acceptor (A_0) are coordinated by one polypeptide, while the accessory BChl is coordinated by the other polypeptide of the dimer.

The distances between cofactors (**Table A2**) can be used to explain the differences observed in ET rates between PSI and the HbRC. Pump-probe spectroscopic measurements previously estimated the lifetimes of the first two ET steps ($RC^* \rightarrow P_{800}^+A_0^- \rightarrow P_{800}^+F_X^-$) in the HbRC of *H. modesticaldum* to be ~25 ps and ~700 ps, respectively [21,30], in line with measurements in HbRCs from other species [31,32]. The distance between the special pair and primary acceptor is too far to allow charge separation to occur in a single step on the picosecond timescale, necessitating the involvement of the Acc cofactor, which is <3.5 Å from both P_{800} and A_0. Two different

models for charge separation can be proposed: one in which Acc serves as the initial acceptor (HbRC* $\rightarrow P_{800}^{+}Acc^{-} \rightarrow P_{800}^{+}A_0^{-} \rightarrow P_{800}^{+}F_X^{-}$) or the initial donor (HbRC* $\rightarrow Acc^{+}A_0^{-} \rightarrow P_{800}^{+}A_0^{-} \rightarrow P_{800}^{+}F_X^{-}$). Future studies informed by this structure will allow testing of these two models.

The role of the menaquinone (MQ) in the HbRC has been controversial. The HbRC that was crystallized in this study lacked MQ (**Figure A4**), as it is not bound tightly by the protein and thus is not present in this structure. However, the edge-to-edge distance between A_0 and F_X is ~10.2 Å (**Table A2**), which is much shorter than the distance between A_0 and F_X in PSI (~14.3 Å) and close enough to allow a maximal ET rate of $\sim 8 \times 10^{8}$ s^{-1}, which is close to the experimentally determined rate of $\sim 1.4 \times 10^{9}$ s^{-1} [21]. Thus, the structure supports the hypothesis that ET from A_0 to F_X in the HbRC does not require an intermediate cofactor. The lack of a quinone as an intermediate cofactor. The lack of a quinone as an intermediate cofactor in forward ET is a striking divergence from other RCs, and the fact that it is not bound tightly by the HbRC may indicate that it can be used as a mobile electron carrier.

Comparison with other Reaction Centers

Because of their descent from a common ancestral protein, all RCs are expected to share the same overall structure [2], but differences in fine detail may lead to distinctions in function. Overall the sequence identity of PshA and the analogous subunits of other RCs is low (<30%), but the structure of the HbRC demonstrates that most of the core structural features are well conserved. **Figures A5 and A6** show a comparison of the TMH arrangement of the antenna domain and ET domain, respectively, of PshA with that

of PsaA and PsaB from PSI, and the CP43 and CP47 chlorophyll-binding subunits of PSII. The antenna domain of PshA (TMH1 to 6) displays structural similarity to the antenna domains of PsaA and PsaB and to CP43 and CP47 (**Figure A5**). This is consistent with the phylogenetic analysis of the antenna domain of Type I RCs and CP43 and CP47, which indicate that the CP43 and CP47 subunits share a common ancestor with the antenna domain of the Type I RC core polypeptide [33]. The conserved structural arrangements of the TMHs did not extend to the surface helices and loops, however, which are considerably different between PSI and the HbRC. This may reflect different interactions with electron donors and acceptors, as well as with PSI subunits that have no analog in the HbRC.

Of the 58 BChls and 2 Chls in the HbRC, 42 of them are in similar positions in PSI (green in **Figure A3C**). The most peripheral antenna pigments from PSI are lacking in the HbRC, probably because PSI uses extra subunits that stabilize their bonding. The most notable differences are in the BChls that link the bulk of the antenna to the ET chain. Three antenna BChl *g* are potentially creating this linkage (teal in **Figure 3.2B and A3A**): (i) the one closest to P_{800} is not found in PSI, (ii) the one closest to Acc is in the same position as an analogous Chl *a* in PSI, and (iii) the one closest to A_0 is shifted toward the P-side by ~5.8 Å relative to a Chl *a* found in a similar orientation in PSI. This makes sense, as the major bridging pigment in PSI is closest to A_0, which is a Chl *a*, like all the Chls of the antenna and ET domains. In the HbRC, energy transfer from an antenna BChl *g* (absorbing at 780-800 nm) to A_0 (8^1-OH-Chl *a* absorbing at 670 nm) would be energetically unfavorable, implying that efficient energy transfer from the antenna to the ET domain should proceed via the Acc or P_{800} cofactors.

The superposition of the HbRC and PSI ET cofactors are displayed in **Figure A7**, and the distances between the HbRC ET chain cofactors are listed in **Table A2**. The distance between Chls or BChls in the special pair is also closer than in other RCs, with a center-to-center distance of only 5.7 Å in the HbRC, compared to 6.3 Å in PSI, 8.2 Å in PSII, and 7.9 Å in the PbRC. The larger overlap of the BChl *g*' macrocycles in P_{800} may explain the high charge delocalization across the BChls in P_{800}^{+} observed by Fourier transform infrared spectroscopy (FTIR) experiments [34]. Both BChl *g*' of P_{800} are coordinated by His sidechains near the beginning of TMH10 (**Figures 3.3 and 3.4A**), which are conserved with PSI. As predicted by FTIR spectroscopy of P_{800}/P_{800}^{+}, the □-N of the His ligand interacts with the Cys601 thiol and the peptide carbonyl of Leu533. In PSI, the special pair (P_{700}) is asymmetric in that (i) the Chl on the PsaA side is Chl *a*', whereas the other is a Chl *a*; and (ii) there is a H-bonding network near the Chl *a*' that is absent on the other side [35,36]. Notably, the protein environment of each BChl *g*' in P_{800} is more similar to that of the Chl *a* on the PsaB side of P_{700} - there is no H-bonding network near the 13^1-keto oxygen, nor are any of the residues involved in the H-bonding network near the Chl *a*' in PSI conserved in PshA. Thus, in the HbRC, the stereochemistry of P_{800} is not correlated with a H-bonding network, leaving the origin and functional importance of this stereochemistry an open question. However, careful examination of the structure suggests that a BChl *g* could not be accommodated in this site, and thus, that the alternate stereochemistry may be driven mainly by steric considerations; the same is true of the lone BChl *g*' in the antenna domain.

The Chl *a* in the Acc position in PSI is not coordinated by the polypeptide. Instead, it is coordinated by a water molecule that is H bonded to an Asn residue [12]. In the

HbRC, the BChl *g* in the Acc position is coordinated by an unknown atom (potentially a chloride ion), which seems to be H bonded to a Gln residue in the same position as the Asn of PSI (**Figures 3.3, 3.4A and 3.4B**). Thus, this cofactor seems to be bound by the protein in a somewhat similar fashion in both PSI and the HbRC.

In PSI, a Tyr from a membrane-parallel helix donates a H-bond to A_0 [12], which is important for tuning the reduction potential of this cofactor [25,37]. This Tyr is replaced in PshA by Ser553, whose hydroxyl group is positioned to donate a H bond to A_0. The shorter distance between the parallel helix and the hydroxyl of the H-bond donor places A_0 closer to the edge of the membrane and thus to F_X. An Arg (PshA-Arg406 from the other subunit) may also be able to H bond to this oxygen; in addition, the proximity of a positive charge could stabilize the A_0^- radical anion during charge separation. This may explain the recent conclusion of Ferlez *et al.* [30] that the reduction potential of the A_0/A_0^- couple is ~150 mV more positive in the HbRC than it is in PSI. It was suggested that this change could also be explained by the replacement of the Met that serves as an unusual axial ligand to the Mg(II) of A_0 in PSI by Ser in PshA. The structure reveals that the axial ligand of A_0 in the HbRC is actually a water that is coordinated by the side chain of Ser545. Both of these differences are expected to influence the reduction potential of the A_0/A_0^- couple.

The quinone-binding site on the A branch of PSI is lined by the indole of PsaA-Trp697 on one side of the napthoquinone headgroup and by the side chains of PsaA-Phe689 and PsaA-Leu722 residue on the other side; the homologous resides are also present in PsaB. Only the Leu is conserved in PshA. In PshA the Phe and Trp are replaced by Met546 and Arg554, respectively; the latter's side chain occupies the space

where the phylloquinone headgroup is located in PSI (**Figure A8A**). Thus, we can say with structural evidence that MQ cannot be bound by the HbRC in the same location as phylloquinone is bound by PSI. Adjacent to A_0 is an extended blob of unassigned electron density, perhaps representing a molecule with a long isoprenoid tail (**Figure A8B**). Although it does not resemble a quinone, it may be in the MQ-binding site; if so, the MQ would not be placed to be an intermediate in ET from A_0 to F_X (**Figure A8C**). In the HbRC, the shorter distance between A_0 and F_X would allow forward ET in the absence of a quinone as an intermediate cofactor.

The F_X cluster-binding site was the only immediately recognizable motif in the PshA sequence when it was first identified [8]. Accordingly, the bonding of F_X to the two conserved Cys in each such motif is similar to F_X binding of PSI (**Figure 3.4C**). However, this cofactor is unusual in that it has a ground spin state of $S = 3/2$ when the cluster is reduced, in contrast to reduced F_X in PSI, which has the usual ground spin state of $S = 1/2$. The symmetric binding site may relieve strain on the cluster, causing this unusual spin state; other known iron-sulfur clusters with a natural ground spin state of $S = 3/2$ include the Fe protein of nitrogenase and 2-hydroxyglutaryl-coenzyme A (CoA) dehydratase, both of which are homodimers with the cluster bound at the dimeric interface [38,39]. The observation that the F_X cluster is such a symmetric cubane that only half of it is seen in the asymmetric unit is consistent with this hypothesis.

In PSI, the PsaC subunit binds two [4Fe-4S] clusters and reduces soluble electron acceptors; it resembles a bacterial-type dicluster ferredoxin and binds tightly to the PsaA-PsaB core because of additional sequence elements [12]. The PshB1 polypeptide, origionally considered the analog of PsaC in the HbRC lacks these elements and is

removed from the $PshA_2$ core early during purification [22], consistent with its absence in the HbRC structure. PshB1 and its homolog, PshB2 [40], are now thought to serve as ferredoxin electron acceptors rather that subunits [41]. An electrostatic surface analysis of the HbRC reveals that the N-side surface is mostly positively charged, with two positive patchesclose to F_X and an extended Lys on either side (**Figure 3.5**). Homology models generated for PshB1 (see supplementary materials) predict a negatively charged surface that may interact with these patches, facilitating reduction by F_X.

Heliobacteria use cytochrome c_{553} (cyt c_{553}) as an electron donor to the HbRC. This is similar in structure to cyt c_6 but is attached to the membrane through covalent linkage to a diacylglycerol [42,43], thus restricting its diffusion to two dimensions. The surface closest to P_{800} is fairly neutral (**Figure 3.5**). There are two surface helices at the P-side of P_{800} that would serve as the possible cyt-binding site. Although similar to that of cyanobacterial PSI [12], the helix in the HbRC is shorter and has minly uncharged sidec hains (Ser, Thr, Asn, Gln) pointing away from P_{800}, and the Trp that is important for cyt binding in PSI [44,45] is absent.

The positions occupied by the PshX subunits are similar to the sites where PsaI (near PsaB) and PsaJ (near PsaA) are found in PSI (**Figure 3.2A**). Both of these are single-TMH subunits, like PshX. PshX and PsaJ both coordinate two antenna BChls, but only one of these is in the same location, and even in this case, the axial ligand is not conserved in the sequence between the HbRC and PSI. PsaI does not coordinate any antenna Chls. Further, there is very low sequence identity shared between PshX and either PsaI (9%) or PsaJ (10%). Thus, it seems likely that the appearance of PshX is a

case of convergent evolution driven by a functional requirement. As with PsaI and PsaJ, a single carotenoid is wedged between the single TMH subunit and the periphery of the core dimer (lime in **Figure 3.1A**). Carotenoids are found in PSI between the TMH of all transmembrane peripheral subunits and the TMH of PsaA-PsaB, implying that, in addition to their functional role, they play a structural role; and this may be the case in the HbRC as well. However, there are only two carotenoids in the HbRC, in contrast to the 22 carotenoids in PSI, and they are shorter as well (30 instead of 40 carbons). This likely reflects the environmental niche of heliobacteria, which are strict anaerobes and therefore do not need to prevent the formation of potentially harmful singlet oxygen.

One might expect the HbRC to display characteristics more similar to the common ancestor of all RCs due to its overall simplicity: its homodimeric core, the low number of antenna BChls relative to its oxygenic homolog, and its lack of peripheral antenna complexes or other subunits (other than PshX) [46]. The conjecture that a homodimeric RC should have preceded a heterodimeric RC in the evolutionary trajectory is also consistent with this idea. However, the HbRC has had a long time to evolve from the ancestral RC and has certainly acquired unique features that are advantageous to the organism. It has been proposed that the Firmicutes, within which the family Heliobacteriaceae reside, branched early in bacterial evolution [7]. Despite this, the ancestor of heliobacteria probably did not originate the first RC, but instead acquired it via horizontal gene transfer, consistent with the colocation of the *pshA* gene in one gene cluster along with all genes required for pigment synthesis [47]. However, some traits of the last common ancestor of all RCs may be preserved in the HbRC as a result of its host's anoxic niche, which has similarities with the early Earth.

Acknowledgements

We would like to extend our gratitude to Julian Whitelegge for his validation of the PshX sequence by mass spectrometry and to Yuval Mazor for modeling and graphical assistance.

This work was funded by the Division of Chemical Sciences, Geosciences, and Biosciences, Office of Basic Energy Sciences, of the U.S. Department of Energy through Grant (DE-SC0010575 to KR, RF, and JHG) and supported with X-ray crystallographic equipment and infrastructure by Biodesign Center for Applied Structural Discovery at Arizona State University.

The Berkeley Center for Structural Biology is supported in part by the National Institutes of Health, National Institute of General Medical Sciences, and the Howard Hughes Medical Institute. The Advanced Light Source is supported by the Director, Office of Science, Office of Basic Energy Sciences, of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231. Results shown in this report are derived from work performed at Argonne National Laboratory, Structural Biology Center at the Advanced Photon Source. Argonne is operated by U. Chicago Argonne, LLC, for the U.S. Department of Energy, Office of Biological and Environmental Research under contract DE-AC02-06CH11357.

Author Information

The HbRC structure has been deposited into the Protein Data Bank with accession code 5V8K. The authors declare no competing financial interests. Correspondence should be addressed to R.F. (Raimund.Fromme@asu.edu).

Figures

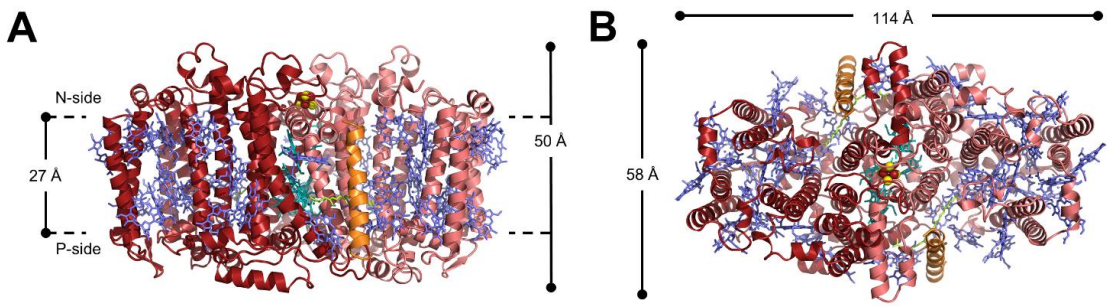


Figure 3.1 Overall structure of the HbRC. Structure as viewed from (**A**) the N-side or (**B**) or within the membrane. The two PshA polypeptides are colored in red and pink. PshX subunits are colored in orange. Cofactor molecules are shown as stick models and colored teal (ET), blue (antenna), and lime (carotenoids). The [4Fe-4S] cluster is shown as red (Fe) and yellow (S) spheres. BChl and Chl tails have been truncated for clarity. Dashed lines denote the approximate hydrophobic boundaries of the membrane bilayer.

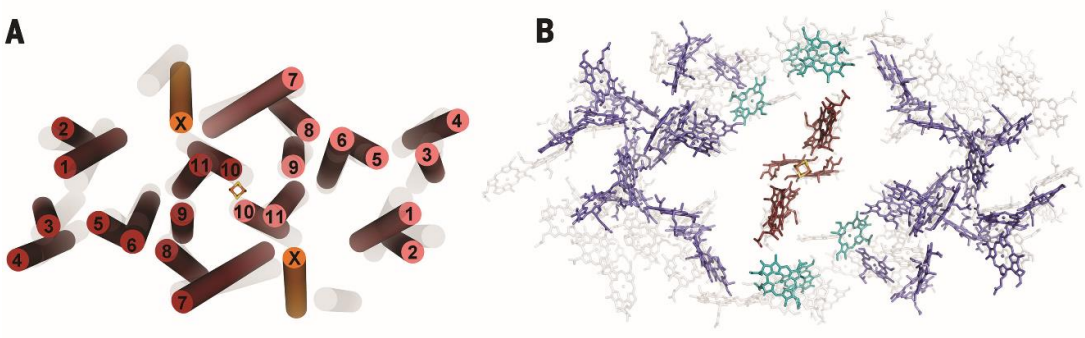


Figure 3.2 Arrangement of the TMHs and pigments in the HbRC compared with PSI. (**A** TMH arrangement of the HbRC and its superposition with the TMHs of PSI (N-side view). TMHs of the two PshA subunits are colored in red and pink and labeled 1-11, from N- to C-terminus. PshX helices are colored in orange and are labeled "X". Transparent grey helices are from the heterodimeric core of PSI [Protein Data Bank (PDB) ID: 1JB0]. (**B**) N-side view of the cofactor organization in the HbRC (colored) superimposed with those associated with the PsaA/PsaB heterodimeric core of PSI (transparent grey, PDB ID: 1JB0). ET (B)Chls are colored red, bulk antenna pigments are colored blue, and the three antenna BChl *g* that flank the ET chain are colored teal. (B)Chl tails have been truncated for clarity.

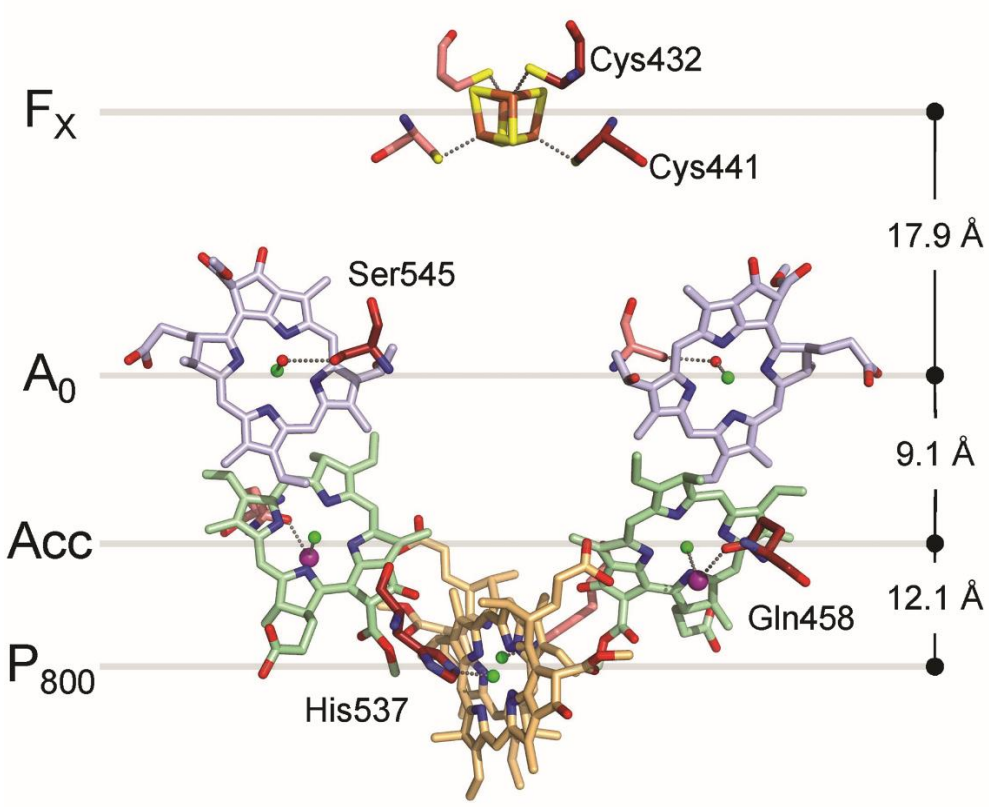


Figure 3.3 Cofactor arrangement of the ET chain. Coordinating residues belonging to one PshA of the homodimer are labeled and colored dark red, while those from the other PshA are labeled and colored pink. Cofactor carbons are colored differently to ease viewing: P_{800} (orange), Acc (green), A_0 (blue). The water molecule serving as an axial ligand to A_0 is shown as a red ball and the unidentified molecule of high electron density ligating Acc is shown as a purple ball. Center-to-center distances are labeled on the right side of the image. BChl and Chl tails have been truncated for clarity.

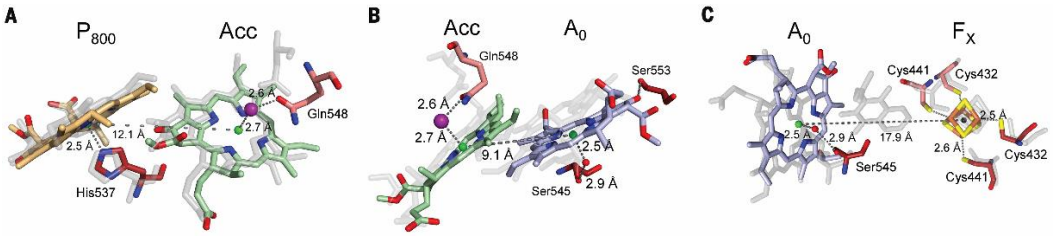


Figure 3.4 Details of the ET cofactors and comparison with PSI. Magnified views of the interactions of PshA with the ET cofactors are shown in panels (**A**) P_{800} and Acc, (**B**) Acc and A_0, and (**C**) A_0 and F_X. Center-to-center distances and ligands are labeled (see also **Table A2**). Coordinating residues belonging to one PshA of the homodimer are labeled and colored dark red, and those from the other PshA are labeled and colored pink. Cofactor carbons are colored to facilitate viewing: P_{800} (orange), Acc (green), and A_0 (blue). The axial ligands to A_0 (water) and Acc (unknown) are shown as a red or purple ball, respectively. The center of the [4Fe-4S] cluster is represented by a black sphere. The corresponding cofactors of PSI are shown in transparent gray. BChl, Chl, and quinone tails have been truncated for clarity.

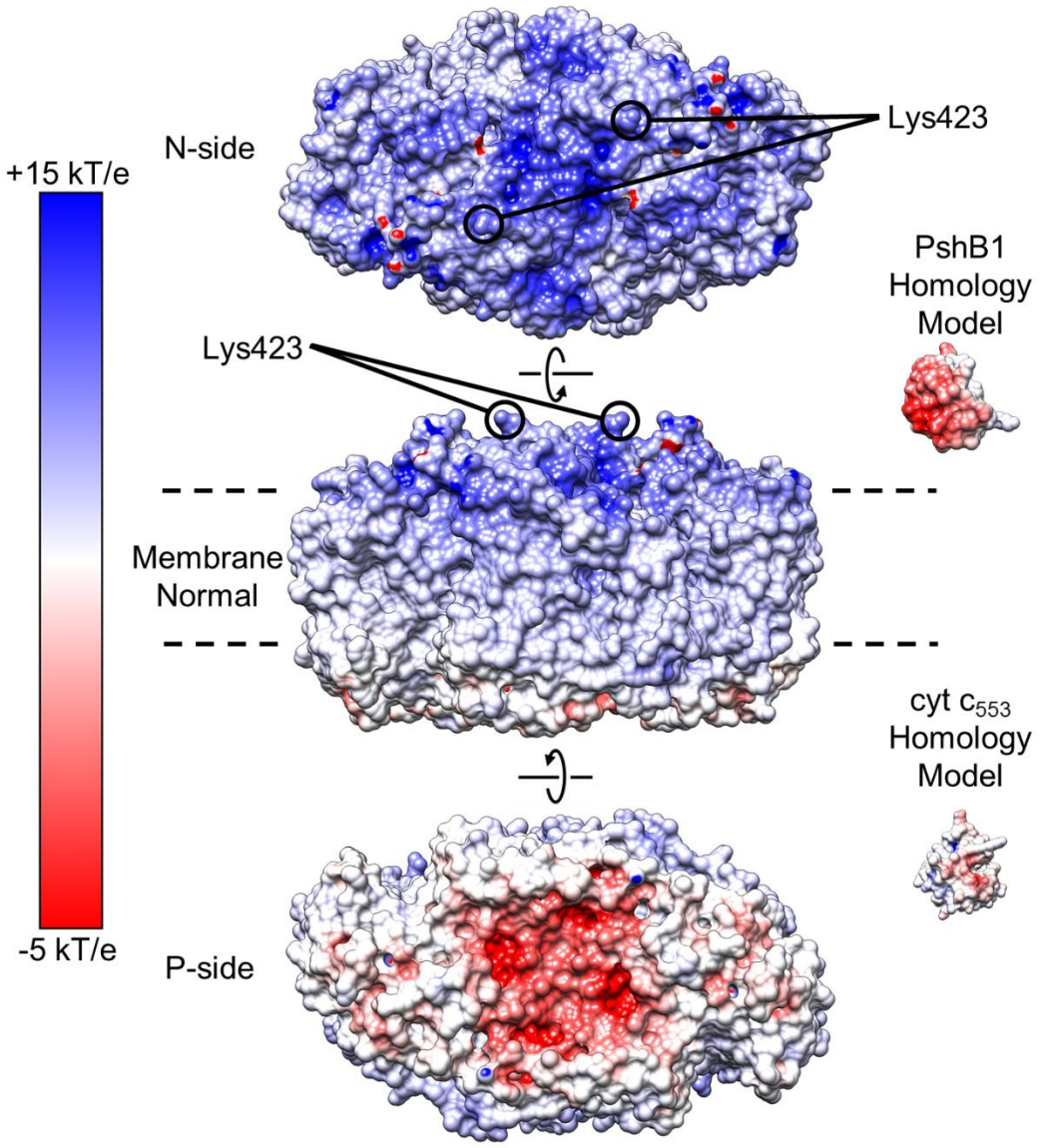


Figure 3.5 Surface electrostatics models of the HbRC. Positive and negative charge are colored in blue and red, respectively (see scale at left). N-side (top) and P-side (bottom) surfaces of the HbRC are shown. The profile view (middle) reveals an asymmetric charg distribution creating an electrostatic field that would stabilize the charge-separated state [48]; membrane boundaries are indicated by dashed lines. The PshB1 ferredoxin and cyt

c$_{553}$ structures shown to the right are multitemplate models based on sequence identity (see Appendix A). Lys584 and Lys587 for a positively charged surface patch on each side of F_X, whereas the side chains of Lys423 (marked) extend out from the HbRC further than any other residue.

References

(1) Schramski, J. R.; Gattie, D. K.; Brown, J. H. *Proc. Natl. Acad. Sci. U. S. A.* **2015**, *112* (31), 9511.

(2) Schubert, W. D.; Klukas, O.; Saenger, W.; Witt, H. T.; Fromme, P.; Krauss, N. *J. Mol. Biol.* **1998**, *280* (2), 297.

(3) Bryant, D. A.; Frigaard, N.-U. U. *Prokaryotic photosynthesis and phototrophy illuminated*; 2006; Vol. 14, pp 488–496.

(4) Gest, H.; Favinger, J. L. *Arch. Microbiol.* **1983**, *136* (1), 11.

(5) Fuller, R. C.; Sprague, S. G.; Gest, H.; Blankenship, R. E. *FEBS Lett.* **1985**, *182* (2), 345.

(6) Trost, J. T.; Blankenship, R. E. *Biochemistry* **1989**, *28* (26), 9898.

(7) Cardona, T. *Photosynthesis Research*. 2015, pp 111–134.

(8) Liebl, U.; Mockensturm-Wilson, M.; Trost, J. T.; Brune, D. C.; Blankenship, R. E.; Vermaas, W. F. J. *Proc. Natl. Acad. Sci. U. S. A.* **1993**, *90* (15), 7124.

(9) Büttner, M.; Xie, D. L.; Nelson, H.; Pinther, W.; Hauska, G.; Nelson, N. *Biochim. Biophys. Acta* **1992**, *1101* (2), 154.

(10) Bryant, D. A.; Costas, A. M. G.; Maresca, J. A.; Chew, A. G. M.; Klatt, C. G.; Bateson, M. M.; Tallon, L. J.; Hostetler, J.; Nelson, W. C.; Heidelberg, J. F.; Ward, D. M. *Science* **2007**, *317* (5837), 523.

(11) Blankenship, R. E. *Photosynth. Res.* **1992**, *33*, 91.

(12) Jordan, P.; Fromme, P.; Witt, H. T.; Klukas, O.; Saenger, W.; Krauss, N. *Nature* **2001**, *411* (6840), 909.

(13) Deisenhofer, J.; Epp, O.; Miki, K.; Huber, R.; Michel, H. *Nature* **1985**, *318* (6047), 618.

(14) Umena, Y.; Kawakami, K.; Shen, J.-R.; Kamiya, N. *Nature* **2011**, *473* (7345), 55.

(15) Kimble, L. K.; Mandelco, L.; Woese, C. R.; Madigan, M. T. *Arch. Microbiol.* **1995**, *163* (4), 259.

(16) Kleinherenbrink, F. A. M.; Ikegami, I.; Hiraishi, A.; Otte, S. C. M.; Amesz, J. *Biochim. Biophys. Acta - Bioenerg.* **1993**, *1142* (1–2), 69.

(17) Olson, J. M.; Blankenship, R. E. In *Photosynthesis Research*; 2004; Vol. 80, pp

373–386.

(18) Khadka, B.; Adeolu, M.; Blankenship, R. E.; Gupta, R. S. *Photosynth. Res.* **2017**, *131* (2), 159.

(19) Lomize, A. L.; Pogozheva, I. D.; Lomize, M. a.; Mosberg, H. I. *Protein Sci.* **2006**, *15* (6), 1318.

(20) Kobayashi, M.; Watanabe, T.; Ikegami, I.; van de Meent, E. J.; Amesz, J. *FEBS Lett.* **1991**, *284* (1), 129.

(21) Chauvet, A.; Sarrou, J.; Lin, S.; Romberger, S. P.; Golbeck, J. H.; Savikhin, S.; Redding, K. E. *Photosynth. Res.* **2013**, *116* (1), 1.

(22) Sarrou, I.; Khan, Z.; Cowgill, J.; Lin, S.; Brune, D.; Romberger, S.; Golbeck, J. H.; Redding, K. E. *Photosynth. Res.* **2012**, *111* (3), 291.

(23) van de Meent, E. J.; Kobayashi, M.; Erkelens, C.; van Veelen, P. A.; Amesz, J.; Watanabe, T. *BBA - Bioenerg.* **1991**, *1058* (3), 356.

(24) Müller, M. G.; Niklas, J.; Lubitz, W.; Holzwarth, A. R. *Biophys. J.* **2003**, *85* (6), 3899.

(25) Müller, M. G.; Slavov, C.; Luthra, R.; Redding, K. E.; Holzwarth, A. R. *Proc. Natl. Acad. Sci. U. S. A.* **2010**, *107* (9), 4123.

(26) Prokhorenko, V. I.; Holzwarth, A. R. *J. Phys. Chem. B* **2000**, *104* (48), 11563.

(27) van Brederode, M. E.; Jones, M. R.; van Mourik, F.; van Stokkum, I. H. M.; van Grondelle, R. *Biochemistry* **1997**, *36* (23), 6855.

(28) van Brederode, M. E.; van Grondelle, R. *FEBS Letters*. 1999, pp 1–7.

(29) Brederode, M. E. van; Mourik, F. van; Stokkum, I. H. M. van; Jones, M. R.; van Brederode, M. E.; van Mourik, F.; van Stokkum, I. H.; Jones, M. R.; van Grondelle, R.; Universiteit, V.; Bank, W.; Kingdom, U. *Proc. Natl. Acad. Sci. U. S. A.* **1999**, *96* (5), 2054.

(30) Ferlez, B.; Cowgill, J.; Dong, W.; Gisriel, C.; Lin, S. S.; Flores, M.; Walters, K.; Cetnar, D.; Redding, K. E. K. E.; Golbeck, J. H. J. H. *Biochemistry* **2016**, *55* (16), 2358.

(31) Nuijs, A. M.; van Dorssen, R. J.; Duysens, L. N. M.; Amesz, J. *Proc. Natl. Acad. Sci. U. S. A.* **1985**, *82* (20), 6865.

(32) Lin, S.; Chiou, H. C.; Kleinherenbrink, F. A.; Blankenship, R. E. *Biophys. J.* **1994**, *66* (2), 437.

(33) Mix, L. J.; Haig, D.; Cavanaugh, C. M. *J. Mol. Evol.* **2005**, *60* (2), 153.

(34) Noguchi, T. *Photosynthesis Research*. 2010, pp 321–331.

(35) Nelson, N.; Yocum, C. F. *Annu. Rev. Plant Biol.* **2006**, *57* (1), 521.

(36) Fromme, R.; Fromme, P.; Grotjohann, I. In *Structural Biology of Membrane Proteins*; Buchanan, S., Grisshammer, R., Eds.; Royal Society of Chemistry, 2006; pp 320–348.

(37) Li, Y.; van der Est, A.; Lucas, M. G.; Ramesh, V. M.; Gu, F.; Petrenko, A.; Lin, S.; Webber, A. N.; Rappaport, F.; Redding, K. *Proc. Natl. Acad. Sci. U. S. A.* **2006**, *103* (7), 2144.

(38) Heinnickel, M.; Agalarov, R.; Svensen, N.; Krebs, C.; Golbeck, J. H.; Park, U. V; Pennsyl, V. *Biochemistry* **2006**, *45* (21), 6756.

(39) Hans, M.; Buckel, W.; Bill, E. *Eur. J. Biochem.* **2000**, *267* (24), 7082.

(40) Romberger, S. P.; Castro, C.; Sun, Y.; Golbeck, J. H. *Photosynth. Res.* **2010**, *104* (2), 293.

(41) Romberger, S. P.; Golbeck, J. H. *Photosynth. Res.* **2012**, *111* (3), 285.

(42) Albert, I.; Rutherford, A. W.; Grav, H.; Kellermann, J.; Michel, H. *Biochemistry* **1998**, *37* (25), 9001.

(43) Kashey, T. S.; Cowgill, J. B.; McConnell, M. D.; Flores, M.; Redding, K. E. *Photosynth. Res.* **2014**, *120* (3), 291.

(44) Sommer, F.; Drepper, F.; Hippler, M. *J. Biol. Chem.* **2002**, *277* (8), 6573.

(45) Sommer, F.; Drepper, F.; Haehnel, W.; Hippler, M. *J. Biol. Chem.* **2004**, *279* (19), 20009.

(46) Gupta, R. S. *Photosynthesis Research*. 2003, pp 173–183.

(47) Sattley, W. M.; Madigan, M. T.; Swingley, W. D.; Cheung, P. C.; Clocksin, K. M.; Conrad, A. L.; Dejesa, L. C.; Honchak, B. M.; Jung, D. O.; Karbach, L. E.; Kurdoglu, A.; Lahiri, S.; Mastrian, S. D.; Page, L. E.; Taylor, H. L.; Wang, Z. T.; Raymond, J.; Chen, M.; Blankenship, R. E.; Touchman, J. W. *J. Bacteriol.* **2008**, *190* (13), 4687.

(48) Gunner, M. R.; Nicholls, A.; Honig, B. *J. Phys. Chem.* **1996**, *100* (10), 4277.

Chapter 4

Evolution of photosynthetic reaction centers: insights from the structure of the heliobacterial reaction center

Christopher Gisriel[1,2,3], Gregory S. Orf [1,2,3], and Kevin E. Redding[1,2]

[1]School of Molecular Sciences, [2]Center for Bioenergy and Photosynthesis, Arizona State University, Tempe, AZ 85287, USA

[3]G.S.O. and C.G. contributed equally to this work.

Abstract

The proliferation of phototrophy within early branching prokaryotes represented a significant step forward in metabolic evolution. All available evidence supports the hypothesis that the photosynthetic reaction center (RC) – the pigment protein complex in which electromagnetic energy (*i.e.* photons of visible or near-infrared light) is converted to chemical energy usable by an organism – arose once in the history of this planet. This event took place over 3 billion years ago and the basic architecture of the RC has diversified into the distinct versions that now exist. Using our recent 2.2-Å X-ray crystal structure of the homodimeric photosynthetic RC from heliobacteria, we have performed a robust comparison of all known RC types with available structural data. These comparisons have allowed us to generate hypotheses about structural and functional aspects of the common ancestors of extant RCs and to expand upon existing evolutionary schemes. Since the heliobacterial RC is homodimeric and loosely binds (and reduces) quinones, we support the view that it retains more ancestral features than its homologs from other groups. In the evolutionary scenario we propose, the ancestral RC predating the division between Type I and Type II RCs was homodimeric, loosely bound two quinones, and performed a slow, light-driven disproportionation reaction to reduce quinone to quinol. The changes leading to the diversification into Type I and Type II RCs were both responses to the need to accelerate this reaction: the Type I lineage added an Fe-S cluster to facilitate double reduction of a quinone, while the Type II lineage heterodimerized and specialized the two cofactor branches, fixing the quinone in the Q_A site. After the Type I/II split, an ancestor to Photosystem I fixed its quinone sites and heterodimerized as a protective response to rising oxygen conditions after the appearance

of the oxygen-evolving complex in an ancestor of Photosystem II. These pivotal events gave rise to the diversity that we observe today.

Introduction

The first bacteria to harness photosynthesis 3.0-3.5 Gya began the process of transforming the Earth's environment into one capable of sustaining multicellular life through their manipulation of natural metabolic gradients and atmospheric composition, especially after the evolution of the water splitting reaction [1,2]. The complex unique to photosynthesis, which facilitates endergonic electron transfer (ET) using light energy, is the reaction center (RC). Light is first absorbed by antenna pigments and the excitation energy is transferred into the center of the RC which houses ET cofactors. Upon reaching these cofactors, the excitation energy triggers a primary charge separation (CS) event, and an electron is transferred down a potential gradient formed by the chain of ET cofactors to a terminal acceptor [3,4]. The CS event is fundamentally important; it is the moment in which the energy contained within a molecular electronic excited state is converted into the energy of a new, biologically-useful redox state. In general, RCs are classified by their terminal electron acceptor. The terminal acceptor of Type I RCs is a [4Fe-4S] cluster, which typically reduces soluble ferredoxin proteins [5]. Type II RCs reduce a mobile quinone to quinol, which then diffuses into the surrounding membrane [6,7].

Seven extant bacterial phyla contain photosynthetic representatives: Acidobacteria, Chlorobi, Chloroflexi, Cyanobacteria, Firmicutes, Gemmatimonadetes, and Proteobacteria [2,8]. Additionally, an ancient symbiosis between a cyanobacterium and

eukaryotic ancestor gave rise to the eukaryotic algae, and from one of these arose the land plants [9]. Due to billions of years of evolutionary divergence between photosynthetic organisms and their wide distribution across different branches of the tree of life, the sequences of the core RC polypeptides have also widely diverged [8]. Their low sequence homology has made piecing together the evolutionary history of photosynthesis difficult even though the number of available genomes for photosynthetic organisms has greatly grown in recent years. However, despite the low sequence homology, fully-folded RC polypeptides retain a surprising degree of structural similarity [10].

All known RCs contain the common structural motif of a dimer of five transmembrane helices (TMHs) with the ET cofactors sandwiched at the dimer's interface, referred to here as the "ET domain". All known Type II RCs exist as heterodimers of 5-THM proteins. In the anoxygenic Type II RCs from Chloroflexi, Proteobacteria, and Gemmatimodetes, these are called the L and M (also known as PufL and PufM, respectively) polypeptides [11,12], while in the oxygen-evolving Photosystem II from Cyanobacteria and eukaryotes, these are called D1 and D2 (also known as PsbA and PsbD, respectively) [13]. The core of the Type I RC from Cyanobacteria and eukaryotes, called Photosystem I (PSI), is a heterodimer of the PsaA and PsaB polypeptides [14,15], while the core of the Type I RCs from the anoxygenic Chlorobi, Acidobacteria, and Heliobacteria are homodimers of the PscA or PshA polypeptides [12,16–18]. In Type I RCs an additional six TMH domain that binds antenna pigments, referred to here as the "antenna domain," is fused to the N-terminus of the 5-TMH ET domain. A duplication and diversification event of a 6-TMH antenna-binding motif gave rise to the modern CP43 and CP47 (also known as PsbC and PsbD, respectively) proteins that bind antenna

pigments and associate with D1 and D2 of Photosystem II, functioning analogously to the antenna domain of the Type I core RC proteins [14,19,20].

Several important questions remain concerning how the RC polypeptides radiated amongst the various prokaryotic lineages, as well as the sequence of events that lead to the distinction in function between Type I and II RCs. Because RC core polypeptides from different reaction centers share low sequence identity (<25%), improving the amount of structural data available for RCs, both in resolution and diversity of phototrophic groups represented, is vital in addressing these evolutionary questions. When sequence identity falls below ~25%, colloquially termed the "twilight zone," evolutionary relationships become too difficult to reliably map [21,22]. Assigning these relationships between groups of photosynthetic bacteria and their proteins is also complicated by the lack of consensus regarding the evolutionary relationships between bacterial phyla in general [23]. However, structural similarity can be used to infer evolutionary relationships well into the twilight zone, resolving discrepancies in sequence alignments leading to more accurate construction of phylogenetic trees [8,10,21,24]. In this review, we have focused on the evolutionary relationships between the RC polypeptides, leaving open the questions regarding the radiation of bacterial lineages.

Recently, the X-ray crystal structure of the Type I heliobacterial RC (HbRC) from *Heliobacterium modesticaldum* was solved at 2.2-Å resolution and deposited in the Protein Data Bank (PDB) under code 5V8K [25]. The HbRC crystal structure revealed a dimer of the PshA polypeptide with two copies of a novel single-TMH antenna subunit PshX bound at the periphery of PshA, in complete C_2 symmetry. This structure is the first representative of a homodimeric RC. Because homodimeric RCs are believed to be the

evolutionary predecessors of heterodimeric RCs [26,27], this structure is an important milestone in the study of the evolution of these proteins. A general scheme of the HbRC and the membrane in which it resides is presented in **Figure 4.1**. Its Gram-positive host [28] lacks internal membranes or invaginations [29], so the HbRC lies within the single membrane that surrounds the cell, inside of the peptidoglycan cell wall. This membrane architecture is unlike chloroplast-containing organisms, where membrane proteins involved in photosynthesis are contained in an extensive network of thylakoid membranes (see review [30]). Upon a light-absorption event, the HbRC transfers an electron through its ET cofactors, referred to as P_{800} (primary electron donor), ec2 (analogous to the term "accessory" used elsewhere), ec3 (primary electron acceptor, analogous to the term "A_0" used elsewhere), and F_X (a [4Fe-4S] cluster), across the membrane reducing a soluble ferredoxin on the acceptor side (**Figure 4.1**). The now-oxidized HbRC is re-reduced on the donor side by a membrane-anchored cytochrome (cyt c_{553}) [31,32].

In this review, we use previously-discussed biophysical considerations in conjunction with structural comparisons to generate hypotheses about the defining characteristics of two important RC ancestors: the last common Type I ancestor of the HbRC and PSI, and the last common ancestor of all photosynthetic RCs. We use new sequence alignment strategies and biological comparisons to consider the sequence of events leading to the diversity of modern RCs. Finally, through the lens of our mechanistic analysis, we propose an evolutionary scheme to explain the Type I/II split and the various heterodimerization events that have occurred since.

What Did the Last Common Ancestor of Type I Reaction Centers Look Like?

Comparison between the new HbRC X-ray crystal structure and PSI is necessary to identify unifying characteristics and ancestral features of Type I RCs. Homodimeric RCs are thought to be the evolutionary precursors to heterodimeric RCs, as the homodimeric state is a simpler state. Duplication of the gene encoding the RC polypeptide followed by divergence of the two genes would result in a heterodimeric RC if the heterodimeric state possessed advantageous properties that were selective [26]. Therefore, the specific differences between the HbRC and PSI caused by the heterodimerization event are important. In this section, we compare the HbRC and PSI, piece-by-piece, considering (1) how the heterodimerization event resulting in PSI affected the ET chain and (2) electron donors/acceptors in each system, (3) how the distance between the ET cofactors affects function, (4) the role of the quinones, and (5) the role and evolution of small extrinsic TMH subunits.

Functional differences between a homodimeric and heterodimeric RC. Functional analyses of extant homodimeric and heterodimeric RCs is required to determine their evolutionary relationships. The heterodimeric PSI has been well-characterized, both functionally and structurally, for many years. In PSI, there are many consequences of the core being heterodimeric, some of which are listed here:

1. P_{700} is a dimer consisting of one Chl *a* bound by PsaB and one Chl *a*' bound by PsaA. This results in a hydrogen-bonding network on the PsaA side that is not present on the PsaB side. It has been hypothesized that this asymmetry localizes

the spin density of the oxidized P_{700}^{+} more on the PsaB-bound Chl *a* [15]. (Note that Chl *a*' and Chl *a* have equivalent reduction potentials.)

2. ET from the phylloquinones on the A- and B-side (PhQ_A and PhQ_B) to F_X exhibit different rates [33].
3. The presence of the asymmetric PsaC, the acceptor-side subunit housing terminal Fe-S clusters F_A and F_B, requires an asymmetric interface for binding [34]. Having a permanently-bound F_A/F_B subunit is thought to reduce the probability of $P^{+}F_X^{-}$ charge recombination (CR) by allowing ET to always progress directly to F_B, increasing the distance between the radical pair in the charge-separated state.
4. In some species, PSI contains a trimer of intimately positioned, π-stacked antenna Chl associated with PsaB; in the analogous position of PsaA, there is a Chl dimer. The trimer has been proposed to be a potential site of low energy-absorbing "red" chlorophyll molecules [15].

Information on the Type I homodimeric RCs remains meager in comparison to PSI. Although the RC from green sulfur bacteria was previously known to have functional similarity to PSI, it was not found to be a homodimer until 1992 [35]. Since then, the other homodimeric RCs that have been discovered (*i.e.* from the heliobacteria and chloracidobacteria) have revealed an important relationship: essentially all of the homodimeric RCs exist within primarily anaerobic groups [36,37]. This suggests that maintaining a homodimeric Type I RC is generally dependent upon a stable anoxic environment.

Association with electron donors and acceptors. Depending on the species, P_{700} of PSI can be reduced by soluble plastocyanin and/or cytochrome c_6 [38]. In contrast, P_{800} of the HbRC has been shown to be reduced by a membrane-anchored cyt c_{553} in vivo (**Figure 4.1**) [31,32]. The membrane-anchored cyt c_{553} reduces P_{800} in 100-700 μs in the HbRC [39,40], compared to the ~3-7 μs reduction of P_{700} by cyt c_6 in PSI in vivo [41,42]. Whether the ancestral Type I RC exhibited RC-bound, membrane-bound, or free electron donors is unknown, but the Gram-positive membrane architecture was probably not present in the Type I RC ancestor. It has been hypothesized that the heliobacterial ancestor gained its RC via lateral gene transfer (LGT) because most genes required for photosynthesis are found in a single cluster in the genome [43], and because the heliobacteria are the only photosynthetic members of the Firmicutes [44]. The HbRC structure from *Hbt. modesticaldum* exhibits an extended loop region between TMH 9 and 10 of the ET domain (**Figure B1**). In all RC structures currently available, this loop contains what we define as the "P-helix", a short surface helix near the special pair that has been hypothesized to be important in PSI for association of electron donors to a site near P_{700} [45]. In the HbRC, this loop is extended; it is ~12 amino acids longer than the analogous loop in PSI, ~26 amino acids longer than the analogous loop in PSII, and ~28 amino acids longer than the analogous loop in the PbRC. It could be the case that the HbRC evolved this extended loop to optimize its interaction with membrane-bound donors, with their trajectories governed by their limited diffusion by virtue of being bound to the membrane. These interactions may allow the cyt c_{553} to better traverse over the surface of the HbRC to reach P_{800}, but future docking simulation studies will likely give insight into the specific interaction with cyt c_{553} and the donor side of the HbRC.

On the acceptor-side, PSI's terminal electron-accepting [4Fe-4S] clusters, F_A and F_B, are contained within the permanently-bound PsaC subunit which extends into the stroma to reduce soluble ferredoxins [15]. The HbRC does not contain an analogous, permanently-bound subunit [46–48]. Therefore, ferredoxins must directly dock near F_X, which lies toward the donor-side of the RC at the interface of the core dimer (**Figure 4.1**). Some of these transiently-binding ferredoxins have been identified and studied [47,49]. Because the heterodimeric core is probably a requirement for the asymmetric PsaC binding in PSI [34,50–52], it is not surprising that the homodimeric HbRC does not exhibit a permanently-bound analogous subunit. If a homodimeric RC preceded a heterodimeric RC, it is likely that ancestor, too, did not exhibit a PsaC-like subunit.

Cofactor distances. A surprising finding in the new HbRC structure was the arrangement of the ET cofactors on the acceptor side. Although the overall arrangement the ET cofactors was very similar of the ET cofactors of PSI, the primary donor (P_{800}) and terminal electron acceptor (F_X) are ~2.5 Å closer together (center-to-center) than the analogous cofactors in PSI. Additionally, ec2 is moved ~2.2 Å closer to F_X and ec3 is moved ~2.4 Å closer to F_X (center-to-center measurement, for a detailed figure, see Gisriel, 2017 [25]). Thus, F_X is closer to the set of 6 chlorins that perform primary CS, and the major difference with PSI is the shorter distance between ec3 and F_X. Another, starker, difference is the lack of a permanently-bound quinone in the HbRC structure (further discussed in the next section), which is unlike any previous RC structure. The shorter distance between ec3 and F_X in the HbRC likely allows forward ET to proceed without the aid of a quinone intermediate, unlike in PSI. This observation begs a question of priority: did the common ancestor of Type I RCs exhibit the shorter distance

arrangement, like the extant HbRC, or the longer distance arrangement, like that exhibited by all other RCs? The observation that only the oxygen-tolerant PSI exhibits this longer distance implies that the presence of oxygen in Earth's atmosphere resulted in the increase in the distance between P and the terminal electron acceptor. The permanent binding of a quinone would allow F_X to reside at a further distance from P_{800}, reducing the probability of CR. This in turn decreases the probability of triplet P_{800} formation (further discussed later), which is beneficial in oxygen-rich conditions since triplet P_{800} can generate singlet oxygen. Again, it is likely that Type I RCs evolved from a homodimeric ancestor [26], which provides support for the Type I ancestor exhibiting the shorter distance arrangement and thus, no permanently-bound quinone intermediate required for ET to F_X.

The role of the quinones. Quinones are used by all photosynthetic RCs, as well as by many other proteins in various bioenergetic pathways. In PSI, the phylloquinones are permanently-bound intermediates in ET between ec3 and F_X. In the Type II RCs (i.e. PSII and the PbRC), the two quinones have distinct roles: one (Q_A) is permanently bound as an intermediate in ET to the other (Q_B), which is the mobile terminal acceptor [53,54]. Two CS events are required to sequentially provide the two electrons to fully reduce Q_B; two protons, which originate from the acceptor-side, must also be provided to Q_B during the process. The second ET event is thought to be a proton-coupled ET ($Q_B^{-\bullet} + e^- + H^+ \rightarrow Q_BH^-$), which is quickly followed by a proton transfer to yield quinol [55]. The quinol then exits the Q_B site and is replaced by a new quinone from the membrane, reinitializing the acceptor side of the RC and increasing the reduction state of the Q-pool [53,54]. The use of

quinones as mobile electron carriers has been one of the defining characteristics of Type II RCs [26].

Experiments probing the involvement of a quinone in the HbRC ET chain have produced conflicting results [49,56–63]. It has been clearly shown that the HbRC does not require quinones for forward ET to the terminal electron acceptor, F_X [63]. However, approximately 4-5 MQ per RC have been observed in membranes of *Hbt. modesticaldum* and *Heliobacillus mobilis*, and menaquinones are always found associated with purified HbRC in variable amounts depending on preparation conditions [25,48,64]. However, the recent HbRC structure does not contain quinones. It was noted, however, that an ill-defined area of electron density lies near ec3 [25]. Although the flat quinone headgroup does not fit into the electron density, an isoprenoid tail with approximately the length of a menaquionone-9 does. One hypothesis to explain this is that during the purification/crystallization process, the loosely-bound quinones exchanged with a geranylgeranyl pyrophosphate, a biosynthesis intermediate of some terpenes and terpenoids, or perhaps a geranylgeranyl phosphate (resulting from hydrolysis of the pyrophosphate moiety) [25]. Interestingly, the location of the electron-dense headgroup of this unassigned density does not lie between ec3 and F_X. Instead, it lies ~5 Å to the periphery of ec3 conjugated macrocycle (*i.e.* farther from F_X than A_0). If this site is where MQ normally binds, then it would be much more consistent with a site for double reduction of a mobile quinone than one for binding a cofactor serving as an ET intermediate between ec3 and F_X.

Recent results demonstrated light-driven quinone reduction in *Hbt. modesticaldum* membranes [65] and strongly suggested that the HbRC is responsible for

this reduction. The data were consistent with a redox cycle between the HbRC and cyt *bc* complex, mediated by MQ (from RC to cyt *bc*) and the membrane-attached cyt c_{553} (from cyt *bc* to RC). Moreover, the inhibitor terbutryn, which binds the Q_B mobile quinone sites of Type II RCs, inhibited the MQ photoreduction activity. If this peripheral site identified in the structure is a mobile quinone reduction site, then the HbRC exhibits the ability to use either soluble ferredoxins or lipophilic quinones as electron acceptors. A potential mechanism is the following: (1) CS in the presence of reduced F_X results in the $P_{800}^{+}ec3^{-}F_X^{-}$ state, (2) ET from $ec3^{-}$ to the nearby MQ results in production of the semiquinone ($P_{800}^{+}MQ^{-}F_X^{-}$), which (3) oxidizes F_X^{-} to yield the fully reduced quinol ($P_{800}^{+}MQH_2$). Protonation of the reduced quinone species must also occur twice, but it is unknown at which stage these proton transfers take place. The HbRC could thus be the first example a Type I RC performing both Type I and Type II functions, a feature reminiscent of hypotheses regarding a "Type 1.5 RC" or the ancestor of all RCs [10,66,67]. This topic is further discussed below as it applies to the ancestor of all RCs ("*Terminal electron acceptors of the ARC and the Type I/II split.*") Since PSI has been shown to be capable of double-reducing plastoquinone in the menD1 mutant [68] and the HbRC appears to also be capable of double-reducing quinones, we hypothesize that the ancestor of Type I RCs did as well.

Extrinsic subunits. Another surprising discovery in the HbRC structure was the presence of two single-TMH subunits, one on each side of the core. Whereas this polypeptide had not previously been documented or annotated as a coding sequence in the genome, clear positive density in the F_O-F_C map indicated its presence. Its primary sequence was identified by crystallography and mass spectrometry [25]. The single-TMH

subunits PsaI and PsaJ are present in PSI in the same general position, but no sequence homology could be seen between these polypeptides and PshX. Interestingly, the gene encoding PshX lies outside of the photosynthetic gene cluster (PGC), which contains *pshA*, the genes for cyt c_{553}, the cyt *bc* complex, and pigment synthesis. This suggests that the gene encoding PshX may not have been transferred with the other genes in the PGC and was an independent evolutionary invention of the heliobacteria. This implies that small transmembrane subunits may be cases of convergent evolution. Its small size (31 residues) makes its *de novo* evolution conceivable. Genomic DNA sequences capable of encoding hydrophobic peptides of sufficient length to cross the membrane are not very rare [69]. If one of these had some affinity for the HbRC and managed to render the complex more stable in some way, then it would be selected for. The same sort of selection may well explain the numerous small hydrophobic subunits at the periphery of PSI, PSII, cyt b_6f, NADH dehydrogenase, etc. Often, deletion of such subunits results in a lower steady-state level of the complex due to more rapid degradation (e.g. PsaJ in eukaryotic PSI [70]). In any case, the PshX subunit found today has two antenna pigment (BChl g) binding sites, and thus contributes to 4 of the 54 antenna chlorins (i.e. 7.4%). Attempts to identify the *pshX* gene in the *Heliorestis convoluta* draft genome have so far been unsuccessful. Thus, either this gene has evolved rapidly during the radiation of the heliobacteria, or *Hrs. convoluta* lacks it. It is even possible that it is specific to *Hbt. modesticaldum*, perhaps as an adaption to growth at higher temperatures. More sequences of *pshX* genes from different heliobacterial species will be required to address these issues. The conclusion that transmembrane proteins may associate with the RC promiscuously, exhibiting convergent evolution, leads us to remain neutral as to whether

the LCA of Type I RCs may have exhibited transmembrane subunits other than the dimeric core.

What Did the Last Common Ancestor of Extant Reaction Centers Look Like?

Sequence alignments of extant RCs: *A phylogenetic approach*. All RCs are related to a distant common ancestor that appeared very early in the history of life on Earth, well before the radiation of the major bacterial phyla observed today [8,10,71–74]. The deconvolution of the evolution of photosynthetic RCs requires a thorough phylogenetic analysis made with reliable sequence alignments, but the low sequence homology between core polypeptides makes this task difficult [8,75]. Despite this, various features of RCs are structurally well-conserved [2,10,73,76]. Therefore, comparison of solved RC structures affords additional data useful for phylogenetic analysis and every additional solved structure increases the confidence in sequence alignments [10] (for a comparison between structure-based and non-structure-based multiple sequence alignments (MSA) from the HbRC and PSI, see **Figures B2** and **B3**). The presence (or lack) of a fused antenna domain also clouds evolutionary analysis. All Type I RCs contain the N-terminal 6 TMH antenna domain fused to the ET domain, but it is found in the form of separate subunits CP43 and CP47 (PsbB and PsbC) in PSII, and is absent in the Type II RCs of the Proteobacteria, Chloroflexi, and Gemmatimonadetes.

The recent HbRC structure has added a new class to the available RC structures for comparison. We have separately performed a phylogenetic analysis of all available RC antenna and ET domain structures to infer evolutionary relationships. There are currently 13 structures available for unique antenna domains and 19 structures available for unique ET domains. Screening a variety of alignment trimming strategies and models,

45 phylogenetic trees were produced for each domain. The evolutionary scenario we propose is based on the relationships inferred by our phylogenetic analysis.

Antenna domain phylogenetic trees. In all antenna domain tree topologies, the antenna domains from PSI and PSII clustered together at opposite ends of an unrooted tree with the HbRC antenna domain branching off in between (**Figure 4.2A**). The non-structure-based antenna domain trees exhibited high variation in their topologies, and low bootstrap values (< 53%), for the arrangement of branching between CP43, CP47, and the HbRC antenna domain (near Point A, **Figure 4.2A**). Conversely, all 15 structure-based trees made from PROMALS3D MSAs exhibited > 85% bootstrap values for an arrangement like that shown in the phylogenetic tree from **Figure 4.2A**. All but one of the 15 structure-based trees made from the PDBeFOLD MSA also gave support for this topology, although confidence was lower; bootstrap values were always > 60%. The strongest-supported trees exhibited a common ancestor between PsaA and PsaB, and a common ancestor between CP43 and CP47 (blue and purple points on **Figure 4.2A**). This is in agreement with previous analyses that the CP43 and CP47 antenna subunits shared a common ancestor, and the PsaA and PsaB antenna domains share a common ancestor [19].

Because all RCs other than the PbRC have either a fused or unfused 6-TMH antenna domain, it is presumed that this domain across all those RCs shares a common ancestor. This hypothesis is supported by the recent finding that superpositions of the antenna domain from the HbRC superimpose with CP43, CP47, and the antenna domains from PSI, with equally low RMSD [25]. The structure-based phylogenetic analysis here also supports this hypothesis, but does not provide clear evidence of one being more closely related to another. However, in our structural comparisons of the antenna domains, we

identified a conserved surface β-hairpin found only in CP43, CP47, and the HbRC antenna domain (see **Figure B1**). This may imply that CP43, CP47, and the HbRC antenna domain share a common ancestor that also exhibited this β-hairpin, the function of which is unknown.

As reasoned previously [19], the placement of the root within any of the main phylogenetic groups would overcomplicate the tree, and require drastically different rates of evolution. Because this is unlikely, the root should be placed within the interior branches: either between the yellow and red spots (marked “A”), or between the red and blue spots (marked “B”) on **Figure 4.2A**. If the root of the tree is near “B” in **Figure 4.2A**, it implies that the HbRC antenna domain shared a common ancestor with CP43 and CP47 more recently than with the PSI antenna domain, and that the PSI antenna domain has diverged sharply from the most ancestral antenna domain. Conversely, in the case that the root is near “A” in **Figure 4.2A**, it implies that the antenna domains of PSI and the HbRC share a more recent common ancestor, but that they have diverged considerably, and that the CP43/47 antenna has convergently evolved to exhibit more similarity with the HbRC antenna domain. The higher degree of structural similarity between the HbRC antenna domain and CP43/CP47 in terms of TMH arrangement, coupled with the conserved surface β-hairpin, makes the latter hypothesis less appealing. Therefore, we conclude that the most likely position for the root of the antenna domain tree would be near "B". This implies that PSI was not the origin of the antenna domains used for PSII. The LCA of CP43 and CP47 thus had a different origin that was more closely related to the HbRC.

ET domain phylogenetic trees. In all non-structure-based phylogenetic trees of the ET domain, the PbRC core subunits (L and M) are a monophyletic group, the PSII core subunits (D1 and D2) are a monophyletic group, and the ET domains of PSI (last 5 TMH of PsaA and PsaB) are a monophyletic group (**Figure 4.2B**). The HbRC shares a common ancestor with the LCA of PSI and the LCA of all Type II RCs, placing its branching point between PSI and Type II RCs. This overall topology is maintained by all 15 non-structure-based ET domain trees that were made, with bootstrap values for the divergence events consistently above 82%.

Like the antenna domain phylogenetic tree, placement of the root within any of the main groups would be unlikely. The root should be placed within the interior branches: either between the red and green points (marked "A"), or between the teal and blue points (marked "B") on **Figure 4.2B**. If the root is nearer to "A", it would imply that a single early split defined the ET domain function. One descendent would have been the ancestor to all Type I RCs, and the other descendent the ancestor to all Type II RCs. If the root is nearer to "B", it would imply that an ancestor of PSI diverged early, and that an ancestor of PSII, PbRC and HbRC split, with the ancestral HbRC gaining a 4Fe-4S cluster (or Type II RCs losing it) and diverging sharply, exhibiting convergent evolution to PSI. This seems unlikely, however, as the ET domains from PSI and the HbRC superimpose with a low RMSD [25], and can shuttle electrons via a terminal 4Fe-4S cluster, implying that they share a common ancestor. Therefore, we favor the former evolutionary scheme.

The phylogenetic tree topology produced by the non-structure-based MSAs described above, however, is poorly supported by those made from the structure-based

MSAs of the ET domains. In 26 of the 30 phylogenetic trees made from structure-based alignments, a low-confidence topology tree is observed for the Type II ancestry. Similar to the non-structure-based phylogenetic trees, the ET domains cluster near one-another for PSII, PbRC, and PSI, with the HbRC sharing a more recent common ancestor with PSI than Type II RCs, but the topology of how L and M diverged, and from where D1 and D2 diverged, is unresolved, exhibiting low bootstrap value (**Figure B4**). Although this low confidence does not refute the evolutionary scheme inferred by the non-structure-based phylogenetic trees, it probably does signify that the confidence of non-structure-based ET domain phylogenetic trees are artificially inflated because of incorrectly-aligned sequences. A limitation of this method, however, is the relatively low number of sequences used in the phylogenetic analysis. It has been shown by multiple groups that the non-structure-based tree topologies probably better reflects the evolutionary relationships between ET domains of RCs [8,10]. Therefore, we favor the topology exhibited by the non-structure-based trees. As more RC structures become available, however, a different picture of RC evolution may arise.

The pigment content of an ancient antenna domain. If the ARC contained the antenna domain, it was either (1) lost by the Type II lineage during (or just after) the Type I/II split, with CP43/CP47 being re-acquired later by the ancestor of PSII, or (2) lost in the PbRC lineage and disconnected via gene fission in ancestral PSII to produce CP43/CP47. Our conclusion that "B" in **Figure 4.2B** is the most likely position for the root of the antenna domain would argue against the latter hypothesis. If the ARC did not contain a fused antenna domain, this feature was likely acquired just after the Type I/Type II split by the Proto-Type I RC (see next sections), and then passed to PSII later

via LGT. We used a comparison of the B(Chl)-binding sites in the various antenna domains to gain insight into their evolutionary relationships.

The TMHs of the antenna domains in the HbRC and PSI exhibit less structural similarity to each other than the TMHs in the ET domain. We investigated whether this lower structural similarity extends to (B)Chl positions as well. We also included cyanobacterial CP43 and CP47 from PSII in the analysis. We overlaid the antenna regions from PshA (PDB code 5V8K, TMHs 1-6), PsaA/B (PDB code 1JB0, TMHs 1-6), and CP43/47 (PDB 3WU2, all 6 TMHs) and identified groups of (B)Chls with conserved positions (**Figure 4.3**). In terms of wide-scale conservation, eleven (B)Chl positions are completely conserved across PshA, PsaA, PsaB, CP43, and CP47 (**Figure 4.3A**), and one extra position is conserved across PshA, PsaA, PsaB, and CP47 (**Figure 4.3G**). These sites roughly form two rings of (B)Chls in the plane of the membrane, creating a core within the antenna domain.

The new HbRC structure confirmed that PshA does not contain a large number of structurally-unique (B)Chl sites at the periphery of the antenna domain like PsaA/PsaB (**Figures 4.3B and 4.3C**) and is far more similar to CP43/CP47 in total antenna site number [25]. It may seem surprising that the HbRC has not increased its antenna BChl number (to the numbers that PSI has shown is possible) to compensate for the lack of large extrinsic antenna [36,48,64,77]. One of the three PshA-unique BChl sites (**Figure 4.3B**) may serve as a weak energy transfer partner for one of the BChls coordinated by PshX, indicating some adaptation to this small antenna subunit. The other two PshA-unique sites are located among the bulk near the outside edge of the antenna domain, distant from the ET domain.

Looking at other incompletely conserved (B)Chl sites, PshA shares two antenna domain (B)Chl sites previously found only in PsaA/PsaB; these are located near antenna (B)Chls coordinated by the ET domain **Figure 4.3D**). Two antenna domain (B)Chl sites previously found only in CP43/CP47 are conserved in PshA, one of which is the closest site by center-to-center distance to a BChl in PshA that is homologous to Chl_D/Chl_Z in PSII (**Figure 4.3E**). Conversely, PsaA/PsaB/CP43/CP47 share two antenna (B)Chl sites that are not found in PshA; these two sites are located at the protein surface on the acceptor side (**Figure 4.3F**). It is tempting to think that these sites are responsible for accepting energy from the extrinsic phycobilisome; however, these sites are still present in plants, which do not have soluble extrinsic antenna. Due to the conservation discrepancies between the proteins, we can only conclude that the 11-12 antenna-(B)Chl "core" of the antenna domain (**Figures 4.3A** and **4.3G**) was probably present in the LCA of Type I RCs. By adding extra sites in strategic locations around this core during diversification, the antenna domain can advance its function to accept energy transfer efficiently from extrinsic antennas, increase its ability to directly absorb photons, or provide more paths for energy transfer to the ET cofactors. But, the observation that the HbRC antenna domain has at least as much structural similarity to CP43/CP47 as it does to the antenna domain of PSI (and is much more similar to CP43/CP47 in terms of bound (B)Chls), demonstrates a complicated evolutionary history. This will be discussed further in the section entitled "*Recruitment of the antenna domain*."

Antenna chlorophylls bound by the ET domain. There are several antenna (B)Chls in the Type I RCs that are coordinated by residues from the ET domain. In the HbRC, 6 BChl *g* molecules are bound in this fashion by each PshA. About twice as many Chl *a* are

bound by the ET domains of PsaA (11) and PsaB (12) in cyanobacterial PSI. In sharp contrast, the PSII core polypeptides (D1 and D2) each bind a single antenna Chl, which is called Chl_Z/Chl_D (or occasionally D1-Chl_Z/D2-Chl_Z) [78]. No antenna sites are found in the structures of the L/M heterodimer of the PbRC [79]; the His residue used to bind Chl_Z/Chl_D is generally mutated to a Phe or Cys. Based upon our sequence alignments, the Chloroflexi RC (CfxRC) and Gemmatimodetes RC (GmRC) are not expected to contain the Chl_D/Chl_Z site either (**Figure B5**). The "red" Chl trimer site in PsaB (and the homologous Chl dimer in PsaA) is not conserved in the HbRC (**Figure B1**). An overlay of the antenna (B)Chls bound by the ET domains is shown in **Figures 4.4A** and **4.4B**.

There are four ET domain-bound antenna (B)Chl sites that are conserved in the HbRC and PSI in terms of sequence (axial ligand residue) and position. One of these sites is analogous to Chl_Z/Chl_D in PSII, in terms of conservation of axial ligand and position, and lies most closely to the ec2 site of the ET chain. There is also a fifth site in which (B)Chls are in a similar position in the structure, but are not sequence-conserved (**Figures 4.4C** and **4.4D**). Three of the five aforementioned sites are situated most closely to the ec2 site of the ET chain, while the other two lie most closely to the ec3 site. We believe energy transfer from the antenna to ec3 to be unlikely in the HbRC because ec3 is a Chl *a* derivative, thus requiring an uphill step from the lower-energy BChl *g* antenna. The same situation is also likely in the GsbRC/CabRC since their antenna are BChl *a* and ec3 is also a Chl *a*.

Since the Chl_Z/Chl_D site is completely conserved across the HbRC, PSI, and PSII, and lies most closely to the ec2 site, a logical conclusion is that it is the bridging site that

energetically connects the antenna domain (or CP43/47) to the ET chain, with energy transfer usually proceeding to the special pair through the ec2 site, which does not require uphill energy transfer. Since none of the anoxygenic Type II RCs (which lack the 6-TMH antenna domain) contain a Chl_D/Chl_Z-type bridging site, but PSII and all known Type I RCs do, a simple evolutionary scenario to explain this is that the Chl_D/Chl_Z site was present in the ARC and pre-dates the Type I/II split. This site may have been important for either aiding in energy transfer from a now-extinct primordial antenna complex or increasing photon absorption by the ET domain. Following this scenario, this antenna site was lost in the last common ancestor of the PbRC/CfxRC/GmRC. The drive to lose this site may have been the recruitment of an ancient LH1-like antenna. In the three-dimensional structures from extant PbRCs, the Chl_D/Chl_Z site occupies the same space as an LH1 polypeptide; inclusion of a BChl at this position would likely clash with LH1 and inhibit its association with the RC (**Figure B5**). Considering that the closest distance between any BChl in LH1 and the ET core in the PbRC is about 11 Å longer than the distance between the Chl_D/Chl_Z site and the ET core, the ancestor of the PbRC may have sacrificed a close distance between the antenna and the ET core to increase the antenna size. The Chl_D/Chl_Z site was probably advantageous for ancestral PSII to keep, if only for using it as an antenna site. This advantage was magnified after the recruitment of ancestral CP43/CP47 and the OEC, to enhance energy transfer efficiency to the ET cofactors [80,81] and act in a photoprotective role in oxygenic conditions [82,83].

Prime (bacterio)chlorophylls in the special pair. The functional role of (B)Chl', which exhibits reversed stereochemistry about the 13^2 carbon in ring E relative to a typical (B)Chl, in the special pair of Type I RCs has been a point of interest. Chl *a*- and

Chl *d*-containing PSI uses a Chl *a*/Chl *a*' or Chl *d*/Chl *d*' heterodimer as P_{700} or P_{740}, respectively [52,84,85]. The HbRC uses a BChl *g*' homodimer as P_{800} and the GsbRC is predicted to use a BChl *a*' homodimer as P_{840} [48,85–87]. Therefore, in Type I RCs, the homodimeric variety employs homodimeric BChl'/BChl' as the special pair, while the heterodimeric variety employs heterodimeric Chl/Chl' as the special pair. Initial high-resolution X-ray crystal structures of PSI revealed that the P_{700} Chl *a*' participates in a hydrogen bonding network. Specifically, these hydrogen bonds are to PsaA-Tyr735 (from the phytol chain ester group), PsaA-Thr743 (from the C-13^1 keto group) and a water molecule (from the C-13^2 ester group). This water molecule is additionally coordinated by 4 residues of PsaA: Ser607, Thr743, Tyr603, and Gly739. Because differences in hydrogen bonding influence the spin density distribution of excited states, it was shown that the majority of spin density for P_{700}^+ would be localized to the non-hydrogen-bonded Chl *a*, perhaps creating a bias for ET toward one branch of the ET chain [15,52]. However, this was cast into doubt after it was shown that both branches of PSI were active in ET [33,88–90].

The situation is different in the HbRC: the two BChl *g*' molecules of P_{800} do not participate in any hydrogen bonding whatsoever. This indicates that the hydrogen-bonding environment around the special pair neither dictates, nor is required for, the presence of the prime epimer. The best explanation remaining for the presence of prime epimer is minimization of steric clash; the (B)Chls' simply fit better into the cavity afforded for the special pair at the Type I RC dimer interface. In the HbRC, if the P_{800} BChl *g*' molecules were replaced with BChl *g* molecules, the 13^2-methoxycarbonyl group would likely sterically clash with residues of TMH 11, particularly Thr598 and Cys601.

This conclusion is reinforced by the observation of a second BChl *g*' per PshA in the HbRC structure, which is found in the antenna domain and is coordinated by His36. No hydrogen bonding is observed between this BChl *g*' and the protein. However, the 13^2-methoxycarbonyl of a normal BChl *g* in this site would likely clash with the neighboring ET-domain antenna BChl *g* that is coordinated by a water molecule and Lys596. Furthermore, if the Chl *a*' of P_{700} in PSI were a Chl *a*, it would likely clash with Phe598 of TMH 9. Therefore, we believe that there is no energetic role for (B)Chl', but only a structural role – and a relatively trivial one, at that.

No enzyme responsible for making the $C13^2$ epimer of any (B)Chl has been identified. An intriguing hypothesis put forward by Webber and Lubitz [91] was that the PSI polypeptide itself catalyzed the conversion of Chl *a* to a Chl *a*', which would then be stabilized by optimal H-bonding to the nearby water molecule coordinated by several PsaA residues (Tyr603, Ser607, Thr743, Gly739). Residue PsaA-Thr739 would play a crucial role in the postulated mechanism, as the H-bond it donates to the 13^1 keto oxygen could stabilize an enolate intermediate after deprotonation of $C13^2$. However, mutation of this Thr to Ala did not result in loss of Chl *a*' [92], casting doubt on this hypothesis. The simplest explanation is that (B)Chl synthesis can result in a small amount of the $C13^2$ epimer. The enzyme responsible for closure of ring E might even be able to interconvert the two isomers at a low rate to allow sufficient production of the prime version. However, when the $C13^2$ epimer is made, selection for it at a specific site would be according to whichever isomer fits best. Although both epimers could compete for any given chlorin site, the low amount of the prime $C13^2$ epimer would result in its losing this

competition except for those sites where the much more abundant (non-prime) epimer could not bind.

With all of these lines of evidence taken together, we believe it likely that the LCA of Type I RCs contained two (B)Chl′ in its special pair. However, we have no reason to believe that this must be extended to the ARC. If the ARC had (B)Chls′ in its special pair, then the Type II lineage lost them. If the ARC had non-prime Chls in the special pair, then the Type I lineage gained them.

The binding orientation of ec2. In all known RCs, the ec2 (B)Chl is the same type as the major (B)Chl of the RC. When the positions of the RC (B)Chls in all the RC structures are compared, the main difference between Type I and II RCs is the orientation of the ec2 cofactor, which is in turn determined by the position of its axial ligand (**Figure 4.5**). In PSII and the PbRC, the surface "P-helix", within the loop between TMH 9 and 10, of D1/D2 and L/M provides the axial ligand for the ec2 cofactors. In L/M, this is a His residue that directly coordinates the central Mg of ec2. In D1/D2, the axial ligand is a water molecule that usually interacts with a Thr side-chain or Ile backbone carbonyl (for a sequence alignment of this region, see **Table B1**). The residue in the analogous position of Type I RCs equivalent is either a Tyr (in the HbRC) or Trp (in PSI). The Trp residue in PSI does not interact with ec2. In the HbRC structure, Tyr510 provides a hydrogen bond to the 13^2-ester carbonyl oxygen of ec2. The axial ligand to ec2 in Type I RCs is instead a water molecule H-bonded to an Gln (HbRC) or Asn (PSI) sidechain found in TMH 9 (**Figures 4.5A** and **4.5B**). This position in the Type II RCs is an Ala or Pro, lacking the ability to coordinate either a (B)Chl central Mg or water (**Figures 4.5C** and **4.5D**).

The different locations of the axial ligand to ec2 – from the 'side' in Type I RCs or from the 'top' in Type II RC – results in a rotation of almost 90° between Type I and II RCs. This is one of the defining differences between the two types of RCs, and is possibly one of the initial changes during the divergence of the two lineages. A possible photophysical explanation for the change arises from the need to optimize ET to the final electron acceptor in each RC. In early Type I RC's that had recently acquired F_X, there would have been selective pressure to optimize ET to F_X. A solution to this would be to stretch the ET cofactors out so that ec3 moves closer to F_X, providing a better driving force for the final ET step. By changing the axial ligand location of ec2, thus rotating it 90°, ec2 would be able to better fit in between P and ec3, so that the rate of ec3 reduction would not suffer. This scenario is supported by the measurement of the center-to-center distances between P and ec3 (**Figure 4.5**). In the HbRC and PSI, these distances are 2-4 Å longer than in the PbRC and PSII. This means that the ec3-to-F_X distance in Type I RCs is 2-4 Å closer than it would be if they had retained the Type II ec2 orientation.

The metallation state of ec3. Every known Type I RC contains a Chl *a* molecule, or derivative thereof, in the ec3 position, regardless of the pigment content of the rest of the RC. Even the green sulfur bacteria, which use millions of BChl *c*, *d*, or *e* molecules per cell to construct their chlorosomes and thousands of BChl *a* molecules per cell to construct their antennas and RCs, still make two Chl *a* molecules per RC to insert specifically into the ec3 site. No reports have thus far been reported of Type I RC variants replacing the Chl *a* at ec3 with another pigment. The fact that the Chl *a* synthesis pathway is preserved even in anaerobic bacteria for such a small proportion of pigments per cell implies that Chl *a* is essential and may have been the original pigment used in the

first Type I RCs. If this is true, then BChls were a later invention by anaerobic bacteria to better adapt to life after the rise of oxygen, which forced their stratification to lower depths of the water column or soil where oxygen was absent and infrared light was more abundant.

The analogous position in Type II RCs is usually occupied by a demetallated version (pheophytin) of the pigment that occupies the special pair position. For example, in the PbRC, if the special pair contains BChl *a*, then the ec3 site will contain BPheo *a*. An exception to this rule is found in PSII from the Chl *d*-producing cyanobacteria *Acaryochloris marina*, which uses Chl *d* in the special pair (P_{680}) but retains Pheo *a* in the ec3 site [93,94]. Mutants of PbRCs have been produced in which the BPheo is replaced by BChl (the resultant BChl site is termed β) [95,96]. When this change occurs, strong energetic mixing results between the $P^{+}ec2^{-}$ and P^{+}/β^{-} states, preventing efficient ET to the quinones and increasing the probability of CR. Therefore, demetalization of ec3 is probably important for maintaining a redox potential higher than that of ec2.

Type II RCs across the Proteobacteria, Chloroflexi, Gemmatimonadetes, and Cyanobacteria/Eukaryotes use different quinone species (*e.g.* ubiquinone, plastoquinone, etc.) as the final electron acceptor and have different electron donors (*e.g.* water or tetraheme cytochromes), necessitating adjustment to the energetics of the ET cofactors. In Type II RCs, the energetic level of the ec3 site scales with that of the special pair; changes in pigments are combined with changes in the redox potentials of the ET chain to adjust the energetics of the entire RC. This would be automatic, as the reduction potential of a pheophytin (Pheo/Pheo$^{\bullet}$) will be ~300 mV more positive than that of the corresponding Chl (Chl/Chl$^{\bullet}$), making the (B)Chl$^{\bullet+}$(B)Pheo$^{\bullet-}$ charge-separated state the

most thermodynamically favorable one of all the potential (B)Chl-based radical pairs. The next cofactor after (B)Pheo is the Q_A quinone, whose reduction potential will always be higher than any (B)Pheo. Thus, ET from (B)Pheo to Q_A *(i.e.* $P^+(B)Pheo^-Q_A \rightarrow P^+(B)PheoQ_A^-$) will always be favorable. In Type II RCs, the main evolutionary invention has been to set up the system such that ET from Q_A to Q_B is favorable, and that protonation and double reduction is allowed only for Q_B. If one assumes that the ARC was a homodimer reducing both quinones as mobile terminal acceptors (see below), then the primary modification would have been to convert a mobile quinone to a fixed quinone serving only as an ET intermediate (*i.e.* Q_A; see discussion below). The use of different (B)Chls and (B)Pheo as RC cofactors, along with changes in the environment of these cofactors would allow them to utilize different electron donors to P as well as acceptors from Pheo and Q_A.

In Type I RCs, the energetic level of ec3 is generally fixed at the level of Chl *a* and thus the P-ec3 energetic difference will vary as the rest of the pigments in the RC are changed. This is likely necessary because the final electron acceptor in all extant Type I RCs is the F_X cluster, which has a much more negative reduction potential than Q_A in any Type II RC (or any mobile quinone). The reduction potential of Chl *a* is lower than Pheo *a* or any BChl. Thus, overall ET from ec3 to F_X will be most favorable if a Chl *a* occupies the ec3 site. This likely explains why the Heliobacteria, Chlorobia, and Chloroacidobacteria bother to make a version of Chl *a*, which they put in no other site than ec3. If the terminal acceptor of the ARC were a mobile quinone, then it is very likely that the ec3 site was originally occupied by a pheophytin. Conversely, if the ARC

contained an F_X site like that of modern Type I RCs, then a Chl *a* would have been favored there. We will argue later that the former hypothesis is much more likely.

Terminal electron acceptors before and after the Type I/II split. Quinones are ubiquitous across all domains of life. They are found in all known photosynthetic RCs, strongly suggesting that the ARC also contained quinones. Reduction of quinones on one side of the membrane and oxidation of quinols on the other side is a major proton pumping mechanism, leading to ATP synthesis. The discovery that quinones can be fully reduced and perform mobile exchange by the HbRC suggests that mobile quinones were an ancestral feature. If electron sources were relatively abundant in the early anoxic earth (e.g. Fe^{2+}, sulfide, H_2), as is currently thought [2], then ATP production by cyclic electron flow (CEF) was likely to be the primary role of the earliest RCs. Therefore, the idea that the ARC evolved early in the history of life as a homodimeric membrane protein whose sole purpose was to reduce quinones to quinols is a reasonable hypothesis. Cytochromes and quinones, as well as a protein serving the role of the cyt *bc* complex, would have predated the RC [97,98], and their function was to link the electron transport pathways of chemo(litho)trophic metabolism to proton pumping.

If we take the converse opinion that the lack of F_X in Type II RCs is a derived trait, we must ask the question: *what was the driving force for loss of F_X?* F_X functions very effectively in all Type I RCs to reduce the ferredoxin pool, which is important for many cellular processes. Moreover, a CEF system in which ferredoxins are reduced by the RC would result in more proton pumping than one in which quinones are reduced by the RC. Ferredoxins can be used to reduce $NAD(P)^+$, and a NAD(P)H dehydrogenase (*i.e.* Complex I) can be used to pass the electrons to the quinone pool, resulting in 1-2

additional protons pumped per electron transferred, depending upon the coupling of proton pumping to electron transport in the ancient Complex I. Thus, if the ARC had an F_X cluster and were part of such a CEF system, to trade that for one in which less ATP was made would represent a significant decrease in fitness.

If the homodimeric ARC did not contain F_X, it faced a limitation in its chemistry. Following CS and subsequent ET to a quinone and reduction of P, the unstable semiquinone radical could have proceeded in three different pathways (**Figure 4.6**):

1) **Protonation of the semiquinone anion and exchange of the neutral semiquinone ($QH^{\bullet}$) with a new quinone:** Semiquinone disproportionation in the membrane (either spontaneous or catalyzed by an unknown enzyme) could lead to production of a quinol ($2\ QH^{\bullet} \rightarrow Q + QH_2$). Alternatively, the semiquinone could have been re-oxidized in the Q_o (Q_P) site of the cyt *bc* complex or analog thereof.
2) **A second excitation event where ET proceeds down the same branch:** After proton-coupled ET to the semiquinone, along with an additional protonation, the mobile quinone will have been converted to a quinol, which can exchange with a new quinone.
3) **A second excitation event where ET proceeds down the opposite branch:** This would lead to production of a semiquinone radical on the other side. Subsequently, a slow proton-coupled radical disproportionation reaction could occur, accompanied by a final protonation of the acceptor. The newly formed quinol could then be replaced by a new quinone.

In all of these scenarios, successful production of quinol could be hindered by an increased probability of CR. In the first scenario, ET down the same branch before quinone exchange could result in CR of the $P^{+}ec3^{-}$ state. In the second scenario, if the quinone had not been protonated before CS (which is unlikely, given the low pK_a of semiquinones), the energy of $ec3^{-}$ might be sufficiently raised to destabilize the charge-separated states involving this cofactor (e.g. $ec2^{+}ec3^{-}$ and $P^{+}ec3^{-}$). This would lead to ET down the opposite branch (scenario 3). In this last scenario, if the two semiquinones are anionic and the disproportionation reaction is slow, then CS on both branches would be inhibited, leading to CR. Thus, while the ARC would be able to convert light energy into chemical energy by oxidizing a high-potential electron donor, like a cytochrome *c*, and reduce quinone to quinol with uptake of protons from the cytosolic (N) side, it would do so rather inefficiently. We hypothesize that the Type I/II split represents two different solutions to improve the efficiency of this reaction.

In what led to the Type I lineage, a [4Fe-4S] cluster was added to facilitate quinone reduction, which requires two electrons. Nature can tune the reduction potential of Fe-S clusters over a wide range (over 1 V of potential) [99] by changing the environment of the cluster. The two reduction potentials associated with quinone reduction (*i.e.* the $Q/QH^{\bullet}$ and the $QH^{\bullet}/QH_2$ couples) can also be tuned by their environment although the first will always be more negative than the second, as the semiquinone is the least stable species. The protonation state of the semiquinone also plays a role, as its pK_a is typically low enough to ensure that the dominant species is the anion ($Q^{\bullet-}$) at neutral pH. The second ET to Q_B in the PbRC is thought to be a proton-coupled ET ($Q^{\bullet-} + e^{-} + H^{+} \rightarrow QH^{-}$), avoiding the unstable $QH^{\bullet}$ and Q^{2-} states [100]. The reduction potential of the ec3 Chl

a is very low (~ -1V), so it should always be favorable for this cofactor to reduce either the quinone or F_X. Depending upon the reduction potential of F_X, ET should be favorable from semiquinone to F_X ($Q^{\bullet-} F_X \rightarrow Q F_X^{-}$) and from F_X to semiquinone ($F_X^{-} Q^{\bullet-} \rightarrow F_X QH_2$). Of course, the details will depend upon the protonation mechanism, and the stability constant of the semiquinone ($K_S = [QH^{\bullet}]^2/\{[Q][QH_2]\}$). However, even if the reduction potential of the F_X/F_X^{-} couple were below that of the $Q/Q^{\bullet-}$ couple, leading to the production of a semiquinone on each side of the RC, F_X could still serve as a catalyst to accelerate the disproportionation reaction ($2\ Q^{\bullet-} + 2H^{+} \rightarrow Q + QH_2$). If the reduction potential of F_X were tuned to be above that of $Q/Q^{\bullet-}$, then the first excitation would lead to reduction of the F_X cluster. The second CS would result in the $P^{\bullet+}Q^{\bullet-}F_X^{-}$ state, and ET from F_X to the semiquinone would produce quinol (*i.e.* the unstable semiquinone oxidizes F_X^{-}). This latter mechanism (**Reaction Scheme 4.1**) is similar to what we have hypothesized for reduction of a mobile quinone in the HbRC [65]. Note that in this mechanism the quinone site does not need to be between ec3 and F_X, as it is found in PSI. The putative quinone site in the HbRC is to the side of ec3 and is slightly further from F_X than ec3 is, and ec3 is closer to F_X than in PSI by ~2.4 Å, suggesting that ec3 directly reduces F_X. Excitation of a RC in the F_X^{-} state could lead to reduction of the quinone next to the ec3 reduced during CS according to the mechanism in the **Reaction Scheme 4.1**. Note that this scheme ignores the protonation steps that are obligatory during reduction of quinone to quinol. Further work will be required to identify protons donors and a pathway from the N side to the site of quinone reduction.

With F_X in place on the cytoplasmic side of the RC with a reduction potential above that of the previous cofactors, the first electron would always end up there. This

situation was one that natural selection could exploit, as any soluble cellular electron carriers with a higher reduction potential could be reduced by F_X^- if they were able to bind to the cytoplasmic face of the RC, even transiently. Indeed, the acceptor side of the HbRC seems to be quite promiscuous, capable of reducing any acceptor with a potential above its own (~ -500 mV) [49], including those not found in heliobacteria, such as cyanobacterial flavodoxin [101]. In all the lineages with Type I RCs, ferredoxins were recruited from the genome to interface with the RC, perhaps binding weakly at first (allowing semiquinone production most of the time), then interacting better and better to take the majority of the electrons from the RC. The lack of homology between the F_A/F_B proteins associated with Type I RCs supports this scenario. There is no universal Type I RC F_A/F_B protein that radiated amongst prokaryotes along with the RC. For example, in heliobacteria, the PshB ferredoxin found to associate with the HbRC is closely related to clostridial ferredoxins and shows little homology, other than the typical ferredoxin Fe-S cluster binding motif, with the F_A/F_B proteins from the Chlorobi or Cyanobacteria. The same situation is true in those phyla – BLAST searches (data not shown) reveal that each F_A/F_B protein is most closely related to other ferredoxins found in the non-photosynthetic members of their phylum, *i.e.* the Ignavibacteria and Melainabacteria. In the case of PSI, and perhaps the GsbRC as well, this ferredoxin has been modified to bind the RC core tightly as a subunit. Even in this case, the electron acceptor is still a ferredoxin. However, the ferredoxin that serves as an electron acceptor for PSI in cyanobacteria (as well as algae and plants) is completely unrelated to the other F_A/F_B proteins, having only a single [2Fe-2S] cluster.

As noted above, an expanded electron transport cycle that included a Complex I (NADH dehydrogenase) in addition to Complex III (cyt *bc*) would substantially increase the number of protons pumped per electron transferred. This would provide the evolutionary driving force to switch F_X from a cofactor enabling quinone reduction to one whose primary function was to reduce ferredoxins. However, in situations when soluble electron acceptors are in short supply (*e.g.* when the ferredoxin pool is largely reduced), then the Type I RC could fall back on quinone reduction as a back-up system to make a smaller amount of ATP via a short cycle involving only Complex III. This seems to be what heliobacteria do now [65].

In what led to the Type II lineage, heterodimerization allowed for specialization of the two branches. In one branch (A-side), the quinone was immobilized and cut off from a proton source, to convert it from a terminal acceptor into an ET intermediate (Q_A). In the other branch (B-side), the reduction potentials of the ec2 and ec3 cofactors were lowered to inhibit CR between the semiquinone and oxidized special pair ($P^+Q_B^-$ state) via the intermediate cofactors on the B-branch. The reduction potentials of Q_A and Q_B were also tuned to favor ET from the former to the latter. The longer stability of the mobile semiquinone would allow enough time for reduction of P^+ and a second CS to occur, leading to successful double-reduction of Q_B to quinol. Note that this last step is the semiquinone disproportionation reaction ($Q_A^{\bullet-} Q_B^{\bullet-} \rightarrow Q_A Q_BH_2$), which has now been optimized in the Type II RC by specialization of each branch and associated quinone.

A Proposed Trajectory for Reaction Center Evolution

We have combined our informed hypotheses into a single scheme that we propose for the evolution of photosynthetic RCs (**Figure 4.7**). It is, of course, impossible to know if our proposed evolutionary pathway is a faithful recounting of what occurred ~3 billion years ago. We have followed a few guiding principles to be consistent in our hypotheses:

- No step should result in a loss of fitness. (*i.e.* a potential future gain tomorrow cannot be used as a reason to lose something that works well today)
- Given the choice between two possible ancestors, in the absence of a compelling rationale, the simpler one should be chosen. (*i.e.* simpler versions usually precede more complex versions)

ET domain. As discussed above, the quinone-reduction activity of the HbRC argues that ancestral Type I RCs had this ability, which implies that the ARC reduced mobile quinones as well. Our mechanistic analysis also led us to propose that the ARC was a homodimeric RC that bound two mobile quinones and contained no F_X site. Therefore, the ARC may have reduced quinones using a (slow) semiquinone disproportionation reaction, and the Type I/II split represents two different strategies to accelerate this reaction and prevent CR. We will use the postulated homodimeric quinone-reducing ARC as our starting point. It is important to note that the ARC was likely not the first RC to appear; any RCs existing before the ARC are ignored here.

The simplest assumption is that the ARC contained a (B)Pheo in the ec3 position to stabilize the initial radical pair (P^+(B)$Pheo^-$) and that the orientation of ec2 was the

same as in extant Type II RCs. This ec2-ec3 configuration likely favors efficient ET to the quinone, as it does to Q_A in extant Type II RCs. Each branch would have performed ET to its quinone rather independently, as discussed above. Full quinone reduction would have relied upon disproportionation either in the membrane or in the protein, or via ET from ec3 to the semiquinone (which may have been strongly disfavored).

Type I lineage. The branch point in RC evolution from the ARC to the Type I RC lineage was the acquisition of the F_X cluster. In our view, this was the key innovation from which most of the other changes flowed. The position of F_X near the cytoplasmic side would have allowed the pre-existing Fe-S cluster insertion machinery to interact with the site after insertion of the core polypeptides into the membrane and dimerization, which likely explains its position. The homodimeric nature of the RC also means that only two (rather than four) mutations to cysteine would be required in the cytoplasmic loop between the second and third TMH of the ET domain. Thus, this evolutionary step would not have been very difficult. As discussed above, the role of F_X would have been to serve as an ET cofactor to the semiquinone on one side, driving its full reduction to a quinol.

Initially, F_X was serving only to facilitate quinol production. However, once its potential was tuned to optimize reduction of ferredoxins, and the cell began to gain an advantage in ATP production from the longer ET cycle, there would have been a selective advantage in optimizing ET to F_X. One of the ways in which the RC was modified was to use a Chl *a* (or derivative) as ec3, which would have produced a larger driving force for ET to F_X, thus conserving more energy to reduce ferredoxins, which have much lower reduction potentials than quinones. The other modification was to

change the orientation of ec2. The effect of this was to stretch the 6-chlorin system across the membrane, placing ec3 closer to F_X for direct reduction. Thus, both the ec3-F_X distance and the driving force would have been changed to optimize ET from ec3 to F_X. Note too that the change in ec2 orientation would have resulted in changes in the arrangement of the TMHs of the ET domain, with the result that only the prime stereoisomer version of the pigment in the P site could be accommodated there as discussed above. However, this change would have little effect on the photochemistry. All extant Type I RCs have inherited all of these changes, so we term this ancestral form Proto-RC1, from which evolved the modern homodimeric RCs, and PSI.

Type II lineage. Molecular clock studies have recently shown that the D1/D2 ancestor and the L/M ancestor diverged from each other very early after the origin of photosynthesis, with the D1/D2 heterodimerization event and the L/M heterodimerization event occurring soon after the split [20,102]. The signature event in the split between D1/D2 and L/M seems to be the loss of the Chl_D/Chl_Z site in the PbRC lineage. As explained above, the driving force for this loss may be the association of the Ancestral PbRC with an LH1-like antenna complex. In the case of both Ancestral PSII and Ancestral PbRC, the next step in the evolutionary scheme is convergent heterodimerization into D1/D2 and L/M. Again, the purpose of this event was to permanently fix one of the two quinones and alter the RC cofactor environments such that CR is disfavored in the branch with the mobile quinone, resulting in biasing ET to the branch containing the fixed quinone. Further development of this arrangement transforms the ARC from performing random ET to one of two identical quinones to favoring ET to a unique quinone. Once the initial changes had been made (e.g. lowering the midpoint potential of $ec3_B$ and blocking the

quinone entry channel to the Q_A site), there would have been additional benefit in introducing more asymmetry into the RC (e.g. lowering midpoint potential of $ec2_B$ and blocking the proton channel to the Q_A site). Additionally, CR becomes less probable because the ET cofactors in the branch containing the mobile quinone are inactive. Also, because CR is now in the seconds timescale for $P^+Q_A^-$, even if CS occurred in a RC with an empty Q_B site, there would be adequate time for a new quinone to arrive via diffusion before CR occurred.

From the points marked "Ancestral PbRC" on our evolutionary scheme, few changes need to occur to result in the modern PbRC/CfxRC/GmRC lineage. From Ancestral PSII, three changes are required to result in modern PSII, all relating to the antenna domain (CP43/CP47) and the OEC (discussed below). These changes must have occurred before 2.0-2.5 Gya to precede the rise of oxygen levels in the atmosphere [2].

The effects of the rise of oxygen. The rise of oxygen forced anaerobic organisms containing a homodimeric Type I RC to avoid aerobic conditions and had an especially profound effect on the evolution of PSI. Almost every ET-domain modification that occurred on the way from the Proto-RC1 to PSI can be explained by the avoidance of singlet oxygen production resulting from the reaction of O_2 with the ^3P triplet state, which is generated by CR of P^+ec3^-. Thus, the major changes can be explained as an effort to minimize the P^+ec3^- CR reaction. First, the quinones in the Proto-RC1 were immobilized and moved further into the interior of the RC to serve as an intermediate in ET between ec3 and F_X, losing mobile quinone reduction permanently in the process. One might ask why this was necessary, if the Proto-RC1 could already carry out efficient ET from ec3 to F_X in the sub-ns timescale, as it does in the HbRC now [49,103]. One answer

lies in the fact that the F_X^- state can be easily generated even when the electron acceptor pool is limiting. Quinone reduction helps to reoxidize F_X^-, but it results in a transiently empty quinone site; an additional CS event on a branch without a quinone would result in P^+ec3^- CR if F_X was already reduced. The subtler answer is that ET from ec3 to F_X ($\tau \approx$ 0.8 ns) is only ~20-fold faster than CR of P^+ec3^-; this ratio of forward ET is lower than any other ET step in Type I RCs, and means that CR would occur in a small but significant fraction of CS events even if F_X were oxidized. The solution to this conundrum was to move the quinone further into the complex and lose the protonation channels so that it cannot be double reduced and/or leave the site. Now the electron is transferred rapidly from $ec3^-$ to the quinone, due to the short distance and very strong driving force. In fact, the system loses so much energy that the next step (ET from Q to F_X) in PSI is almost isoenergetic or slightly uphill [89].

This seems counter-intuitive in terms of optimizing the rate of ET; the overall rate of ET from P to F_X is much slower in PSI (tens-to-hundreds of ns) than it is in the HbRC (< 1 ns). However, the point was not to increase the overall rate of ET to F_X, it was to minimize P^+ec3^- CR. It has been shown that CR of the P^+Q^- state in PSI does not produce a $^3P_{700}$ triplet [104], thus, having the electron on the quinone is "safe". The net electron throughput of the RC is not determined by the rate of ET to F_X anyway, as the limiting rates involve diffusion of electron donors and acceptor to and from the RC. This evolutionary state, in which the Proto-RC1 has immobilized its quinones and made them permanent members of the ET chain, is Ancestral PSI.

As the residues in the two phylloquinone-binding sites in PSI are highly conserved between the PsaA and PsaB subunits, the immobilization of the quinone sites

must have occurred in the homodimeric state, before diversification of the RC core polypeptide into PsaA and PsaB. It has been argued that the asymmetry of the quinone sites in PSI is due to removing the risk of $^1O_2$ production: as the semiquinone in PhQ_A site is lower in energy than in the PhQ_B site, CR from the FeS clusters would precede through PhQ_A, subsequently tunneling to the ground state [105]. From our argument above, it seems likely that in the homodimeric state the quinone would have been more like PhQ_A (*i.e.* lower in energy in the semiquinone state). After formation of the heterodimer, one of the quinones was free to raise its energy, although the reason for this is presently unclear.

Ancestral PSI would have dealt with another issue related to the presence of atmospheric oxygen: long-lived F_X^- states can generate superoxide (O_2^-), another reactive oxygen species (ROS). However, unlike singlet oxygen, a biological remediation pathway does exist to eliminate superoxide in the form of the enzyme superoxide dismutase. The product of this enzyme is hydrogen peroxide, which can be reduced to water by peroxidases. In effect, Ancestral PSI sacrificed its ability to reduce quinones to prevent formation of an ROS the cell could not remediate (singlet oxygen), in the process allowing occasional formation of an ROS it could remediate (superoxide).

A way to further spatially separate the charge separated states, protecting the RC further, was to recruit a modified version of the ferredoxin acceptor as a permanently bound subunit. In modern PSI, this is the PsaC subunit, which contains the F_A/F_B clusters and is bound to the PsaA/PsaB core. As one would expect, the crystal structures of PSI reveal that PsaC binds asymmetrically to the PsaA/PsaB heterodimer. Therefore, heterodimerization likely occurred in Ancestral PSI to encourage stronger binding to

PsaC, resulting in modern PSI. The result of this would be to further separate the CS state, increasing its lifetime to ~100 ms. Without the F_A/F_B clusters, if the acceptor pool were heavily reduced, the $PQ^-F_X^-$ state could eventually accumulate under high light, leading to detrimental CR of P^+ec3^-. With the additional two clusters, PSI would have to accumulate 4 electrons – on all 3 Fe-S clusters as well as one of the quinones – before P^+ec3^- CR would occur, making this situation very unlikely. It should be noted that this can occur under anaerobic conditions, but the lack of O_2 removes the danger of singlet oxygen, thus explaining why an ancestor to the HbRC did not fix their quinone or heterodimerize; they were not forced to acquire a mechanism to avoid ROS.

Recruitment of the antenna domain. One can imagine that the ARC had an antenna domain. This was then subsequently lost in the PbRC lineage and split off into a separate subunit in the Ancestral PSII lineage. The latter change is a relatively minor one, but the former is a problem. What possible advantage would the ancestor to the PbRC have gained by losing so many antenna pigments? One cannot use the argument that this allowed LH1 antenna complex to bind to it, as LHI would not have been "waiting" for the antenna domain to be lost so it could associate with the ET domain. As mentioned before, we will avoid any step that results in a loss of fitness unless there is a compelling argument for it. There is no such argument here. It seems much more likely that the ARC lacked an antenna domain and the PbRC lineage later gained LH1, which then prevented it from using an antenna domain if one were transferred to it via LGT.

Perhaps an ancestor of PSII gained its antenna domain (**Figure 4.7**) via LGT from a homodimeric ancestor of a Type I RC that contained two copies of a single antenna protein per homodimeric RC. Such an event must have occurred before the Ancestral

PSII gained an OEC (*i.e.* before oxygen was present in the atmosphere) because asymmetry is a requirement for binding of the OEC in modern PSII. From there, the unfused antenna of Ancestral PSII would have later diverged into CP43 and CP47. If the Type I ancestor serving as the donor of the antenna domain was a common ancestor of all Type I RCs, it would imply that both the HbRC antenna domain and CP43/47 remained relatively unchanged since the LGT event, while PSI diverged more, a scenario that we cannot explain. If, however, the Type I ancestor that transferred the antenna domain gene did so after the Type I homodimer/Ancestral PSI split, it would provide a good explanation as to why the HbRC and PSII antenna are so similar: the LGT would have been a more recent event, allowing for less divergence. This would explain the presence of a β-hairpin that is conserved between the HbRC, CP43, and CP47, on the donor side of the antenna domains (**Figure B1C**), although the function of this β-hairpin is unclear.

An alternative evolutionary scheme can be proposed where the ancestral antenna joined and fused to the homodimeric Type I ancestor via LGT from a homodimeric ancestor of PSII that had two copies of an antenna protein symmetrically associated with its homodimeric core. We find this scenario less compelling based on two observations. First, all extant Type I RCs have an antenna domain and only PSII has an antenna domain. Therefore, the LGT would have had to occur from an Ancestral PSII before the Type I lineages split, since the antenna domain is conserved across all extant Type I RCs. As explained above, the fact that the HbRC antenna and CP43/CP47 are so similar cannot be easily explained by this scenario. Second, both the 6-TMH antenna domain and the 5-TMH ET domain of Type I RCs coordinate antenna (B)Chls. PSII has only a single antenna chlorophyll (Chl_D/Chl_Z) coordinated by the ET domain, and the PbRC has none.

The much larger number of antenna sites in the Type I ET domains (*i.e.* 6-12) implies the antenna domain was present in this lineage for longer, enough time for the Type I RCs to add more bridging sites between ET and antenna domains. If PSII was the progenitor of the antenna domain, one would expect that it would have evolved more ET domain-bound antenna sites to aid in EET to the ET core.

Therefore, according to our evolutionary scheme, an antenna domain was gained by an ancestor of all modern Type I RCs, producing what we term Proto-RC1'. After the split the lineages that led to modern homodimeric RCs and PSI, an ancestor to modern homodimeric RCs laterally transferred this antenna domain to an ancestor of PSII.

Conclusion and Outlook

The recent 2.2-Å crystal structure of the HbRC has revealed important insight into the evolutionary trajectories of RCs, which we have noted here. Our scheme rests upon the hypothesis that the last common ancestor of all RCs was a homodimeric complex that functioned to reduce quinone to quinol using a slow radical disproportionation reaction, allowing for cyclic electron flow to be light-driven for the first time. All of the diversification presently observed across the disparate phototrophic taxa are attributable to the need to, first, optimize this reaction, and second, adjust to rising oxygen concentrations. We acknowledge that trying to understand protein-level changes that occurred over three billion years ago is a difficult task. However, we believe that our approach, which considers every *functional* aspect of the RC, can help fill in informational gaps that genetic and structural data alone cannot fill. Recent advances in structural biology methods allow for optimism that representative RC structures from

more phototrophic groups will eventually be solved, and recent advances in metagenomics analysis allow for optimism that new interesting phototrophs will be discovered. Marrying these future findings with our underlying functional analysis will help to further refine our evolutionary picture.

Methods

Sequence and structural alignments. RC structures were gathered from the protein data bank (PDB) [106]. Domains of the polypeptide of interest (either the ET domain or an antenna domain of each RC) was manually extracted from each full structure using the PyMol software [107] and its sequence was fetched from the Uniprot database [108]. **Tables B2 and B3** provide a list of all structures used for the structure-based phylogenetic analysis for antenna and ET domains, respectively. The antenna domain data set included 13 structures and the ET domain set included 19 structures. When comparing RC polypeptides which do not have available structures (such as the CfxRC), these sequences were fetched from the Uniprot database. Two structure-based MSAs were constructed for each dataset using the PROMALS3D server [109] and the Protein Data Bank in Europe PDBeFold server [110]. We also constructed a sequence-based MSA from each data set using Clustal Omega implemented in MEGA7 [111,112]. This resulted in a total of 6 MSAs (two structure-based and one sequence-based for both ET and antenna domain data sets). When showing the three-dimensional overlay of RC structures, the cealign or super function of Pymol was used.

MSA trimming and model selection. We applied 5 different MSA data trimming strategies to each MSA: no gap deletion, 25% site coverage cutoff, 50% site coverage

cutoff, 75% site coverage cutoff, and full deletion for positions containing gaps. To determine which model should be used to assess the evolutionary relationships between the non-structure-based MSAs (Clustal Omega), and structure-based MSAs (PROMALS3D and PDBeFOLD), the "Find Best Protein Model" function of the MEGA7 software [113], which ranks 56 protein substitution models according to Baysian and Akaike information criteria and the Maximum Likelihood fit, was used. The top 5 models for each MSA are shown in **Table B4**. The three most re-occurring models for each data set were identified with the top-scoring specification (with or without gamma distributed rate categories, and with or without proportions of invariant sites or amino acid frequencies from the data). For the antenna domain, the three amino acid substitution models that were used to create phylogenetic trees were cpREV [114], WAG [115], and LG [116]. For the ET domain, the three amino acid substitution models that were used to create phylogenetic trees were LG [116], rtREV [117], and WAG [115]. Trees were built under each of the top three models, for a total of 45 phylogenetic trees for each domain (Table S5). Support for nodes in preliminary trees was assessed by 50 bootstrap replicates.

The resulting 45 tree topologies for each domain were examined for consensus and a tree that exhibited the most common and well-supported topology was chosen to represent the overall data and re-created with 500 bootstrap replicates (**Figure 4.2**). Both trees were made with the following specifications: evolutionary relationships were deduced using the Maximum Likelihood method based on the LG model [116]. The tree exhibiting the highest log likelihood value (-4300.51 for the antenna domain shown in **Figure 4.2A** and -3535.73 for the ET domain shown in **Figure 4.2B**) is shown in **Figure**

4.2. A pairwise distance matrix was estimated with a JTT model and the Neighbor-Join and BioNJ algorithms were applied to identify the initial tree for the heuristic search. The tree exhibiting the highest log likelihood value was selected and a discrete Gamma distribution was used to model evolutionary rate differences among sites (4 categories (+*G*, parameter = 6.3855 for the antenna domain shown in **Figure 4.2A** and 13.1963 for the ET domain shown in **Figure 4.2B**)).

Acknowledgements

The authors thank Gillian Gile (Arizona State University) for consultation on phylogenetic tree construction. All authors are supported by the Division of Chemical Sciences, Geosciences, and Biosciences, Office of Basic Energy Sciences, of the U.S. Department of Energy through grant DE-SC0010575 to K.E.R.

Schemes

Scheme 4.1 Proposed mechanism of quinone reduction by a prototypical Type I RC

First charge separation	Donor reduces P^+	Second charge separation	F_X^- reduces semiquinone

$$Pec3QF_X \rightarrow P^+ec3QF_X^- \rightarrow Pec3QF_X^- \rightarrow P^+ec3Q^-F_X^- \rightarrow P^+ec3Q^{2-}F_X$$

Figures

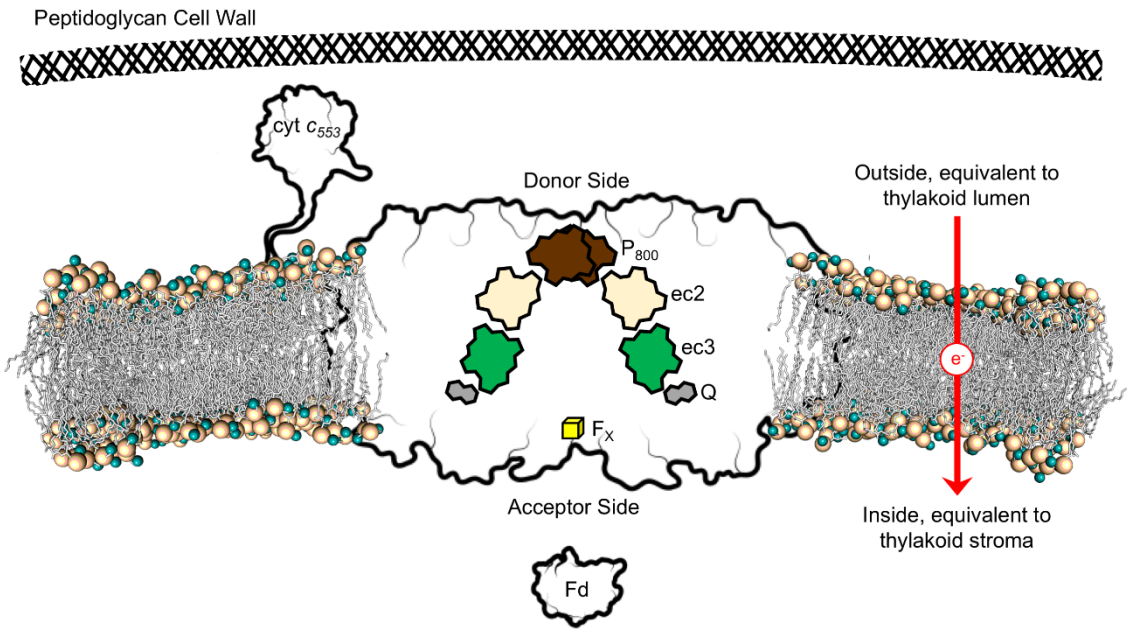


Figure 4.1 Membrane architecture of the heliobacteria as it pertains to the HbRC. An outline of the HbRC crystal structure within the membrane, a cyt c_{553} homology model, and a PshB1 ferredoxin homology model are shown as they are associated with the heliobacterial membrane. The cofactor names and approximate location within the HbRC are shown in boxes. P_{800} is colored in brown, ec2 is colored in tan, ec3 is colored in green, quinones are colored in grey, and F_X is colored in yellow. The general direction of light-driven ET across the membrane is shown by the red arrow on the right.

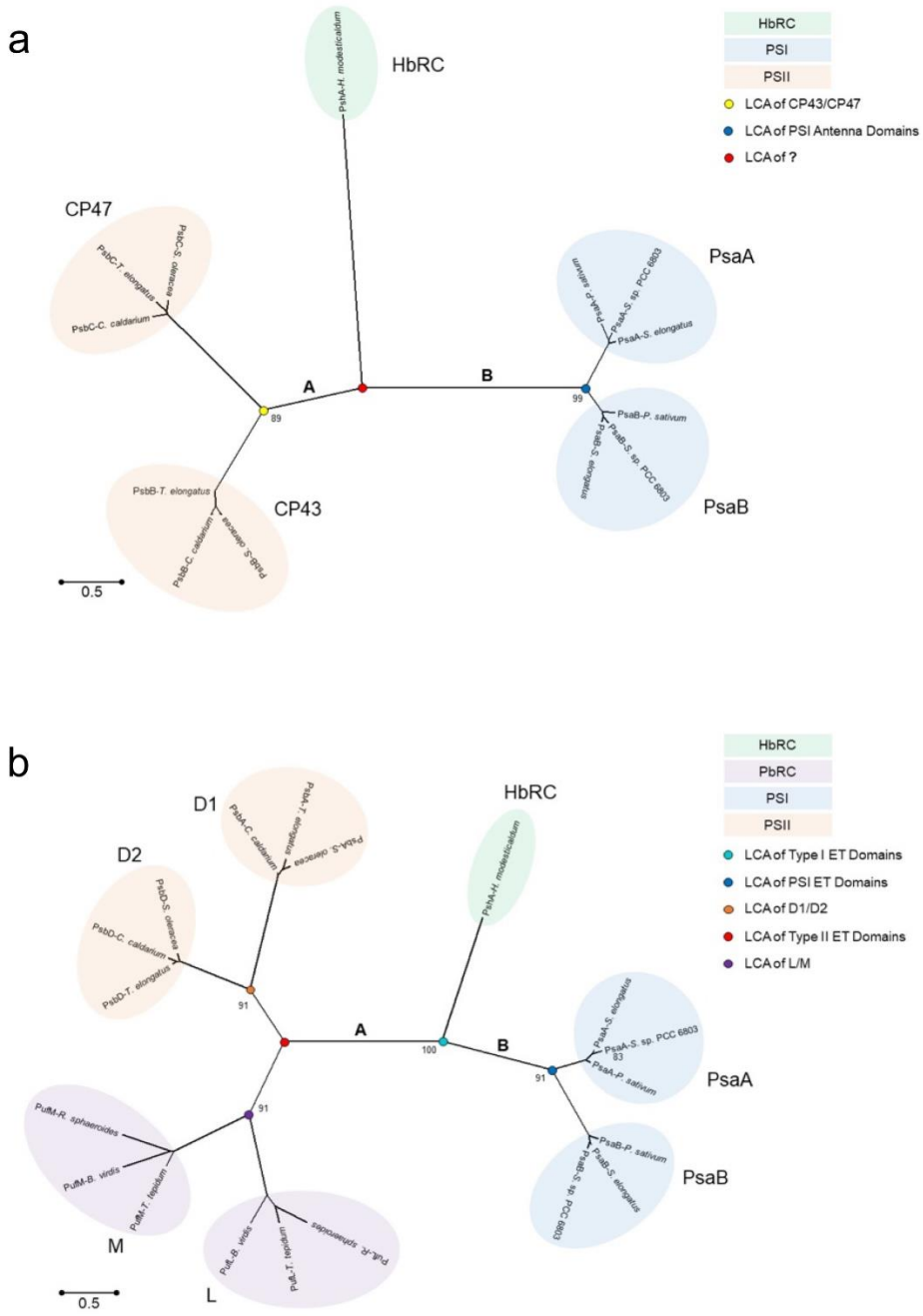


Figure 4.2 Antenna and ET domain phylogenetic trees. Structure-based (PROMALS3D) phylogenetic tree of the antenna domain MSA (**A**) and non-structure-based (Clustal Omega) phylogenetic tree of the ET domain MSA (**B**). The tree is drawn to scale, with

branch lengths measured in the number of substitutions per site (scale bar in bottom left of each panel). The analysis involved 13 amino acid sequences for the antenna domain (**A**) and 19 amino acid sequences for the ET domain (**B**). All positions containing gaps and missing data were eliminated. There was a total of 242 positions in the final dataset for panel A and 135 positions for panel B. The percentage of trees in which in which the associated taxa clustered together in the bootstrap test (500 replicates) are shown next to the branches. The RC with which the polypeptide is affiliated, and the node significance, are listed in the legend. Potential roots are discussed in the text using the letters “A” and “B” for each panel.

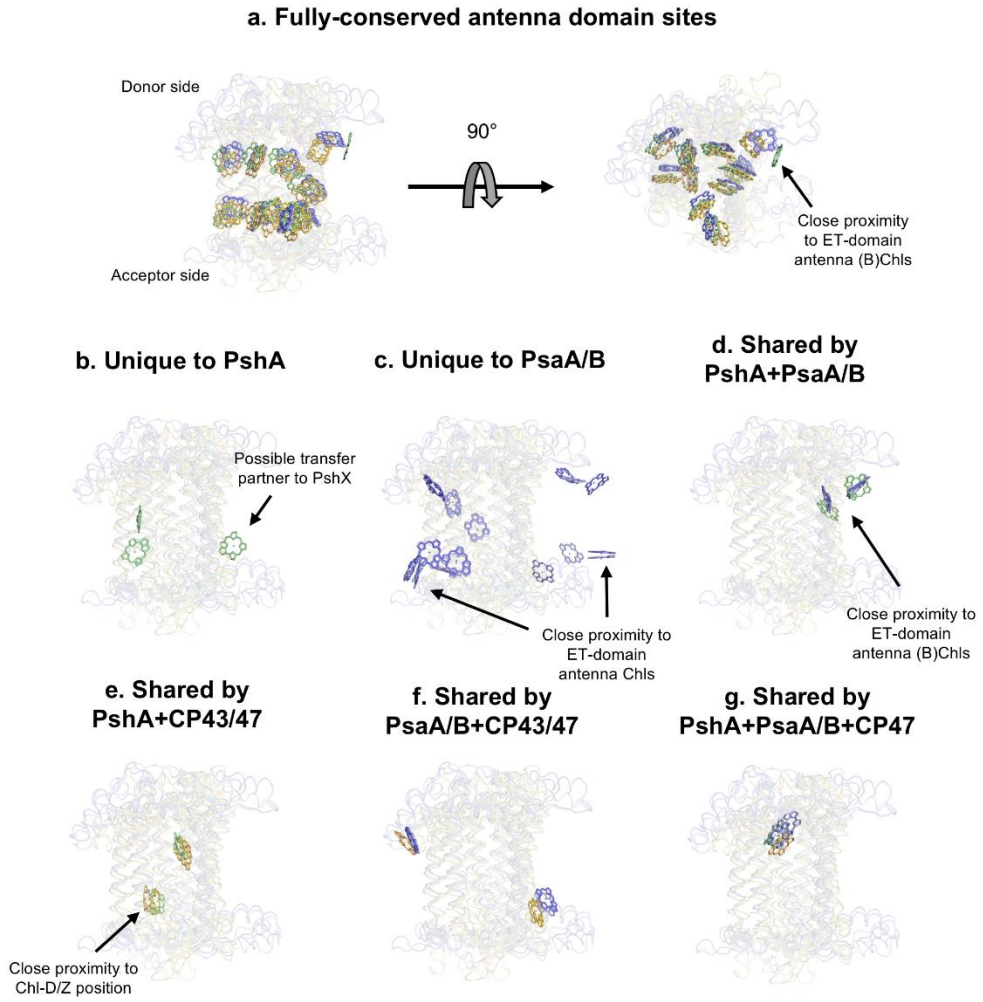


Figure 4.3 Aligned overlays of the core antenna TMHs and pigments from the HbRC, cyanobacterial PSI, and cyanobacterial PSII. The viewpoint is from the distal side of the RC looking inwards toward the ET core with donor side (P-side) up and acceptor side (N-side) down. Polypeptides are shown as transparent ribbons and (B)Chls are shown as macrocycle sticks to facilitate viewing. Green: PshA (PDB code 5V8K); blue: PsaA/PsaB (PDB code 1JB0); yellow: CP43/CP47 (PDB code 3WU2).

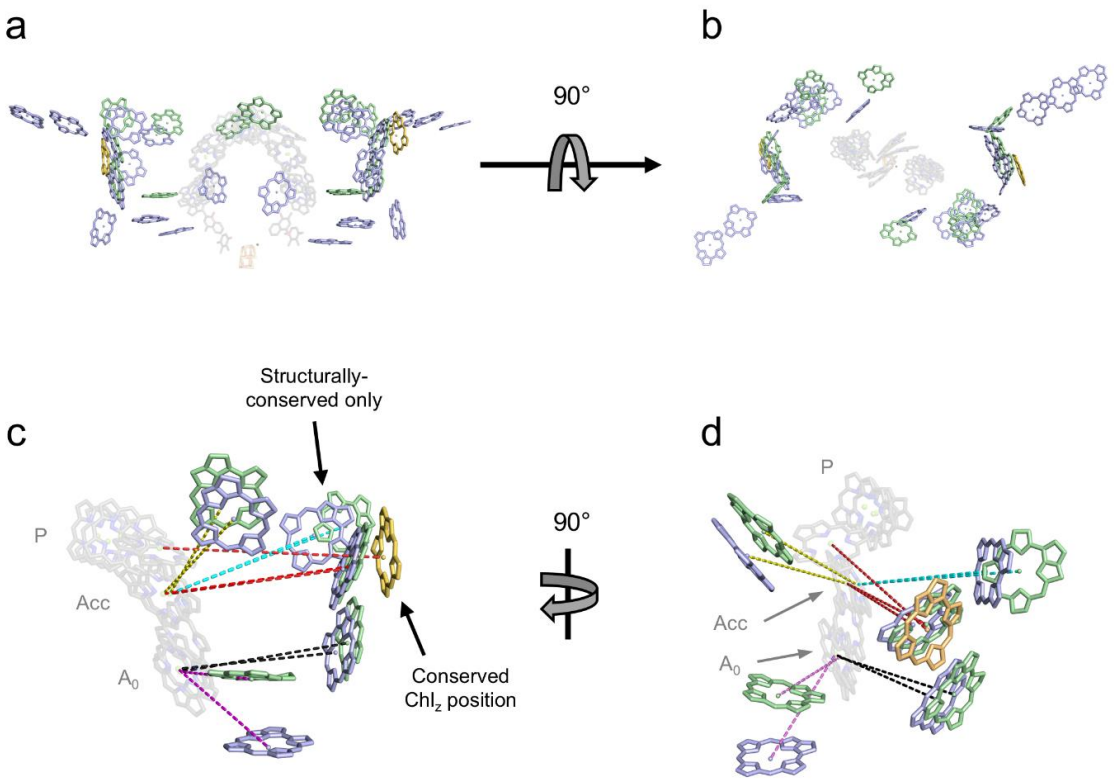


Figure 4.4 Comparison of antenna sites provided by the ET domains of the HbRC, PSI, and PSII. All antenna sites are shown from an in-membrane (**A**) and donor-side (**B**) view. Structurally-conserved sites within 20 Å (center-to-center) of any ET cofactors are shown from a perspective perpendicular (**C**) and parallel (**D**) to the long dimension of the RCs. Gray: ET cofactors; green: HbRC sites; blue: PSI sites; yellow: PSII sites. Pigment macrocycles are shown alone to ease viewing. In (**C**) and (**D**), dotted lines connect the antenna site center-to-center with the closest ET cofactor; lines are color-coded by site.

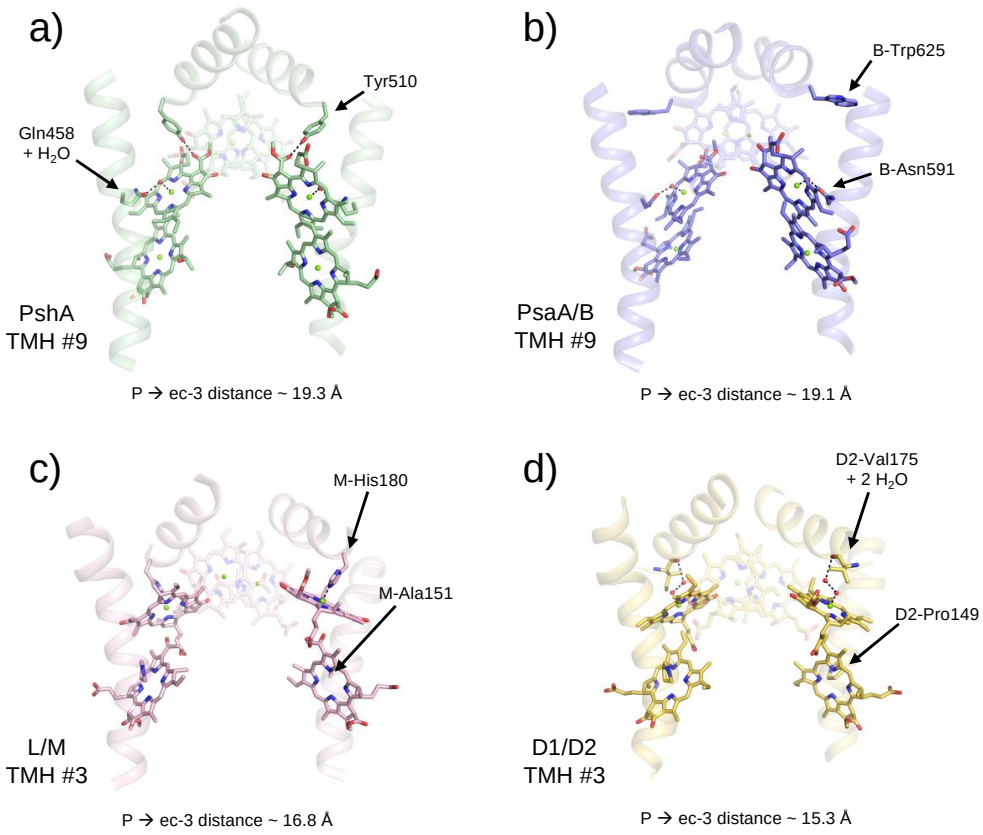


Figure 4.5 Differences in ec2 BChl coordination between the HbRC and PbRC structures. In the HbRC (**A**), Tyr510 descends from the P-helix to provide a hydrogen bond to the 13^1-keto oxygen of the ec2 BChl *g*. The homologous residue in PSI (**B**) is a Trp which does not participate in hydrogen bonding. The homologous residue is a His residue in the PbRC structure (**C**) or an Ile residue plus two waters in PSII (**D**) that provide the axial ligand to the central Mg of the ec2. In (**A**) and (**B**), the axial ligand for the ec2 central Mg is a water molecule coordinated by a Gln or Asn on TMH 9, but in (**C**) and (**D**), this position is an Ala or Pro which does not contribute to ec2 coordination. PDB structures 5V8K, 1JB0 (chain A), 1PRC (chain L), and 3WU2 (chain A) were used to construct this comparison. (B)Chl tails are omitted for clarity. Stated distances between

ET cofactors are center-to-center distances.

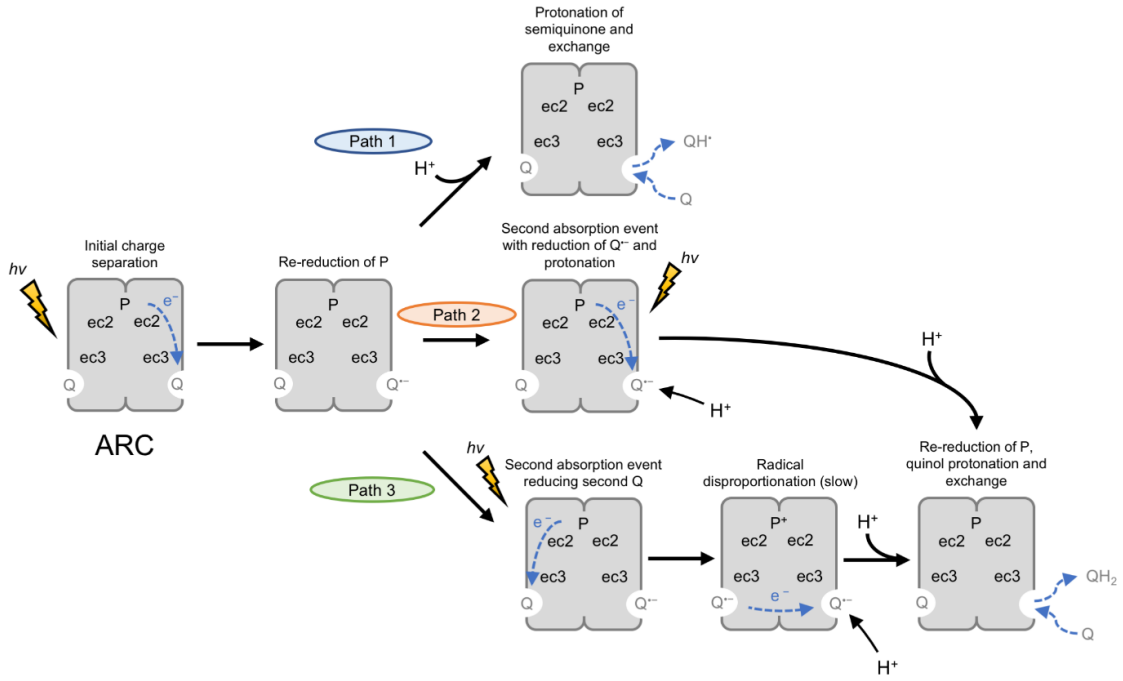


Figure 4.6 Proposed mechanism for quinone reduction in a homodimeric ancestral RC with two mobile quinone sites and no F_X cluster**.** After the initial light-driven reduction of quinone to semiquinone, three paths exist to fully reduce the semiquinone to quinol. At all points before the re-reduction of P^+ by an external reductant, CR is a probable occurrence.

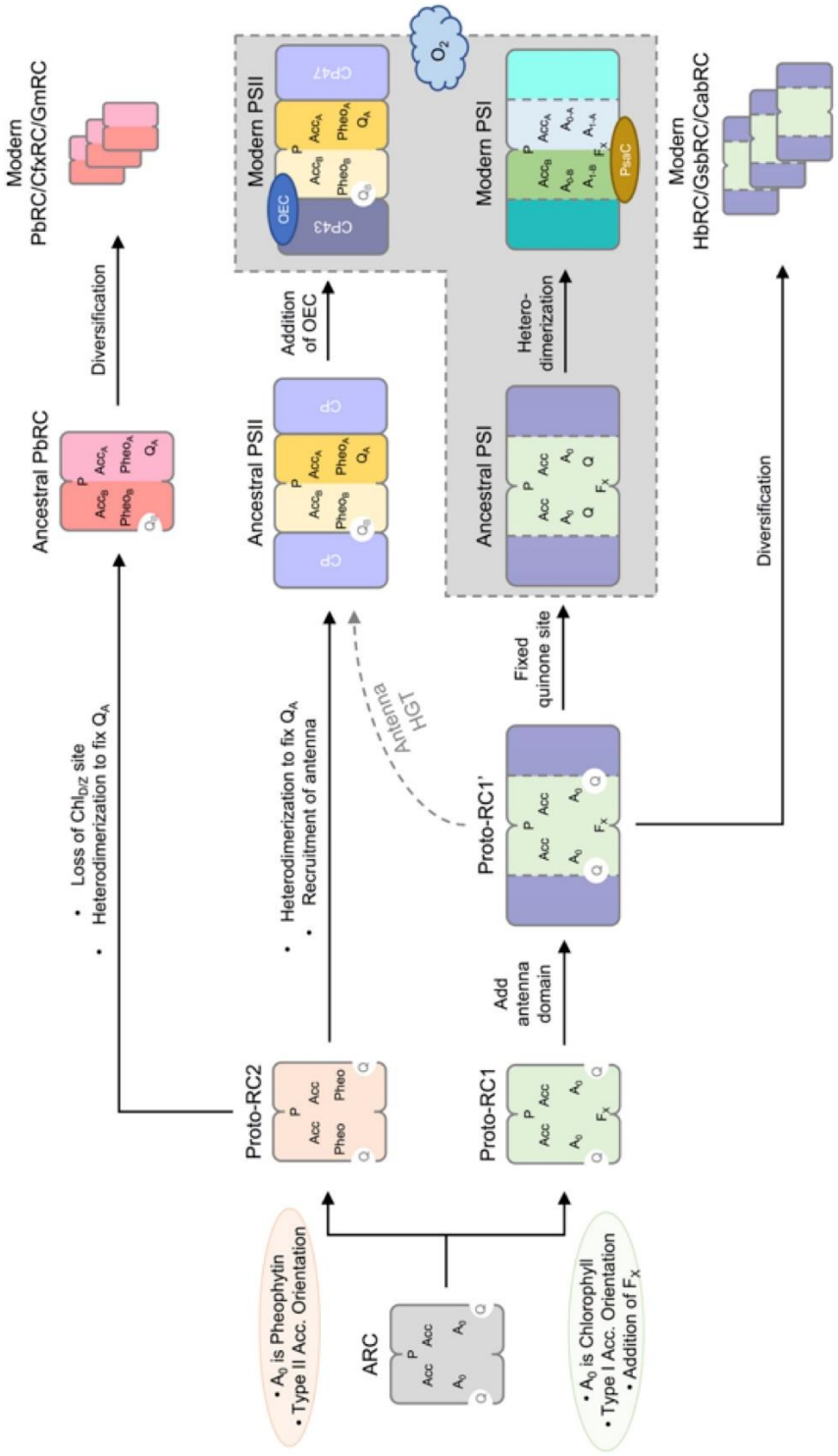


Figure 4.7 Proposed evolutionary scheme for photosynthetic RCs.

References

(1) Hohmann-Marriott, M. F.; Blankenship, R. E. *Annu. Rev. Plant Biol.* **2011**, *62*, 515.

(2) Fischer, W. W.; Hemp, J.; Johnson, J. E. *Annu. Rev. Earth Planet. Sci.* **2016**, *44* (1), 647.

(3) Nelson, N.; Yocum, C. F. *Annu. Rev. Plant Biol.* **2006**, *57* (1), 521.

(4) Blankenship, R. E. *Molecular Mechanisms of Photosynthesis*, 2nd ed.; John Wiley & Sons, Ltd.: Southern Gate, Chicester, West Sussex, UK, 2008.

(5) Vassiliev, I. R.; Antonkine, M. L.; Golbeck, J. H. *Biochim. Biophys. Acta - Bioenerg.* **2001**, *1507* (1–3), 139.

(6) Diner, B. A.; Petrouleas, V.; Wendoloski, J. J. *Physiol. Plant.* **1991**, *81* (3), 423.

(7) Cardona, T.; Sedoud, A.; Cox, N.; Rutherford, A. W. *Biochim. Biophys. Acta - Bioenerg.* **2012**, *1817* (1), 26.

(8) Cardona, T. *Photosynthesis Research*. 2015, pp 111–134.

(9) Keeling, P. J. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* **2010**, *365* (1541), 729.

(10) Sadekar, S.; Raymond, J.; Blankenship, R. E. *Mol. Biol. Evol.* **2006**, *23* (11), 2001.

(11) Zeng, Y.; Feng, F.; Medová, H.; Dean, J.; Koblížek, M. *Proc. Natl. Acad. Sci. U. S. A.* **2014**, *111* (21), 7795.

(12) Bryant, D. A.; Frigaard, N.-U. U. *Prokaryotic photosynthesis and phototrophy illuminated*; 2006; Vol. 14, pp 488–496.

(13) Barber, J. *Cold Spring Harb. Symp. Quant. Biol.* **2012**, *77*, 295.

(14) Schubert, W.-D.; Klukas, O.; Saenger, W.; Witt, H. T.; Fromme, P.; Krauß, N. *J. Mol. Biol.* **1998**, *280* (2), 297.

(15) Jordan, P.; Fromme, P.; Witt, H. T.; Klukas, O.; Saenger, W.; Krauss, N. *Nature* **2001**, *411* (6840), 909.

(16) Rémigy, H.-W.; Stahlberg, H.; Fotiadis, D.; Müller, S. A.; Wolpensinger, B.; Engel, A.; Hauska, G.; Tsiotis, G. *J. Mol. Biol.* **1999**, *290* (4), 851.

(17) Heinnickel, M.; Golbeck, J. H. *Photosynth. Res.* **2007**, *92* (1), 35.

(18) Garcia Costas, A. M.; Liu, Z.; Tomsho, L. P.; Schuster, S. C.; Ward, D. M.;

Bryant, D. A. *Environ. Microbiol.* **2012**, *14* (1), 177.

(19) Mix, L. J.; Haig, D.; Cavanaugh, C. M. *J. Mol. Evol.* **2005**, *60* (2), 153.

(20) Cardona, T. *Front. Plant Sci.* **2016**, *7* (March), 1.

(21) Rost, B. *Protein Eng. Des. Sel.* **1999**, *12* (2), 85.

(22) Doolittle, R. F. *Of urfs and orfs: a primer on how to analyze derived amino acid sequences*; University Science Books: Mill Valley, CA, 1986.

(23) Battistuzzi, F. U.; Hedges, S. B. *Mol. Biol. Evol.* **2009**, *26* (2), 335.

(24) Khadka, B.; Adeolu, M.; Blankenship, R. E.; Gupta, R. S. *Photosynth. Res.* **2017**, *131* (2), 159.

(25) Gisriel, C.; Sarrou, I.; Ferlez, B. H.; Golbeck, J. H.; Redding, K. E.; Fromme, R. *Science* **2017**, *357* (6355), 1021.

(26) Blankenship, R. E. *Photosynth. Res.* **1992**, *33* (October), 91.

(27) Nitschke, W.; Rutherford, A. W. *Trends Biochem. Sci.* **1991**, *16* (7), 241.

(28) Vermaas, W. F. J. *Photosynth. Res.* **1994**, *41* (1), 285.

(29) Miller, K. R.; Jacob, J. S.; Smith, U.; Kolaczkowski, S.; Bowman, M. K. *Arch. Microbiol.* **1986**, *146* (2), 111.

(30) Pribil, M.; Labs, M.; Leister, D. *Journal of Experimental Botany*. 2014, pp 1955–1972.

(31) Prince, R. C.; Gest, H.; Blankenship, R. E. *BBA - Bioenerg.* **1985**, *810* (3), 377.

(32) Albert, I.; Rutherford, A. W.; Grav, H.; Kellermann, J.; Michel, H. *Biochemistry* **1998**, *37* (25), 9001.

(33) Guergova-Kuras, M.; Boudreaux, B.; Joliot, A.; Joliot, P.; Redding, K. E. *Proc. Natl. Acad. Sci. U. S. A.* **2001**, *98* (8), 4437.

(34) Jagannathan, B.; Golbeck, J. H. *Cellular and Molecular Life Sciences*. 2009, pp 1257–1270.

(35) Büttner, M.; Xie, D. L.; Nelson, H.; Pinther, W.; Hauska, G.; Nelson, N. *Biochim. Biophys. Acta* **1992**, *1101* (2), 154.

(36) Liebl, U.; Mockensturm-Wilson, M.; Trost, J. T.; Brune, D. C.; Blankenship, R. E.; Vermaas, W. F. J. *Proc. Natl. Acad. Sci. U. S. A.* **1993**, *90* (15), 7124.

(37) Bryant, D. A.; Garcia Costas, A. M.; Maresca, J. A.; Gomez Maqueo Chew, A.; Klatt, C. G.; Bateson, M. M.; Tallon, L. J.; Hostetler, J.; Nelson, W. C.; Heidelberg, J. F.; Ward, D. M. *Science.* **2007**, *317* (5837), 523.

(38) Chitnis, P. R.; Xu, Q.; Chitnis, V. P.; Nechushtai, R. *Photosynthesis Research*. 1995, pp 23–40.

(39) Vos, M. H.; Klaassen, H. E.; van Gorkom, H. J. *BBA - Bioenerg.* **1989**, *973* (2), 163.

(40) Nitschke, W.; Liebl, U.; Matsuura, K.; Kramer, D. M. *Biochemistry* **1995**, *34* (37), 11831.

(41) Hervás, M.; Navarro, J. a; Díaz, a; Bottin, H.; De la Rosa, M. a. *Biochemistry* **1995**, *34* (36), 11321.

(42) Hippler, M.; Drepper, F.; Farah, J.; Rochaix, J. D. *Biochemistry* **1997**, *36* (21), 6343.

(43) Xiong, J.; Inoue, K.; Bauer, C. E. *Proc. Natl. Acad. Sci. U. S. A.* **1998**, *95* (25), 14851.

(44) Bryantseva, I. A.; Gorlenko, V. M.; Kompantseva, E. I.; Achenbach, L. A.; Madigan, M. T. *Arch. Microbiol.* **1999**, *172* (3), 167.

(45) Fromme, P.; Schubert, W. D.; Krauß, N. *BBA - Bioenergetics*. 1994, pp 99–105.

(46) Heinnickel, M.; Shen, G.; Agalarov, R.; Golbeck, J. H. *Biochemistry* **2005**, *44* (29), 9950.

(47) Heinnickel, M.; Shen, G.; Golbeck, J. H. *Biochemistry* **2007**, *46* (9), 2530.

(48) Sarrou, I.; Khan, Z.; Cowgill, J.; Lin, S.; Brune, D. C.; Romberger, S.; Golbeck, J. H.; Redding, K. E. *Photosynth. Res.* **2012**, *111* (3), 291.

(49) Ferlez, B.; Cowgill, J.; Dong, W.; Gisriel, C.; Lin, S. S.; Flores, M.; Walters, K.; Cetnar, D.; Redding, K. E. K. E.; Golbeck, J. H. J. H. *Biochemistry* **2016**, *55* (16), 2358.

(50) Jagannathan, B.; Golbeck, J. H. *Biochemistry* **2009**, *48* (23), 5405.

(51) Jagannathan, B.; Dekat, S.; Golbeck, J. H.; Lakshmi, K. V. *Biochemistry* **2010**, *49* (11), 2398.

(52) Fromme, P.; Jordan, P.; Krauß, N. *Biochimica et Biophysica Acta - Bioenergetics*. 2001, pp 5–31.

(53) Okamura, M. Y.; Feher, G. *Annu. Rev. Biochem.* **1992**, *61*, 861.

(54) Okamura, M. Y.; Paddock, M. L.; Graige, M. S.; Feher, G. *Biochim. Biophys. Acta* **2000**, *1458* (1), 148.

(55) Graige, M. S.; Paddock, M. L.; Bruce, J. M.; Feher, G.; Okamura, M. Y. *J. Am. Chem. Soc.* **1996**, *118* (38), 9005.

(56) Brok, M.; Vasmel, H.; Horikx, J. T. G.; Hoff, A. J. *FEBS Lett.* **1986**, *194* (2), 322.

(57) Brettel, K.; Leibl, W.; Liebl, U. *Biochimica et Biophysica Acta - Bioenergetics*. 1998, pp 175–181.

(58) Lin, S.; Chiou, H. C.; Blankenship, R. E. *Biochemistry* **1995**, *34* (39), 12761.

(59) van Der Est, A.; Hager-Braun, C.; Leibl, W.; Hauska, G.; Stehlik, D. *Biochim. Biophys. Acta - Bioenerg.* **1998**, *1409* (2), 87.

(60) Trost, J. T.; Brune, D. C.; Blankenship, R. E. *Photosynth. Res.* **1992**, *32* (1), 11.

(61) Miyamoto, R.; Mino, H.; Kondo, T.; Itoh, S.; Oh-oka, H. *Biochemistry* **2008**, *47* (15), 4386.

(62) Kondo, T.; Itoh, S.; Matsuoka, M.; Azai, C.; Oh-Oka, H. *J. Phys. Chem. B* **2015**, *119* (27), 8480.

(63) Kleinherenbrink, F. A. M.; Ikegami, I.; Hiraishi, A.; Otte, S. C. M.; Amesz, J. *Biochim. Biophys. Acta - Bioenerg.* **1993**, *1142* (1–2), 69.

(64) Trost, J. T.; Blankenship, R. E. *Biochemistry* **1989**, *28* (26), 9898.

(65) Kashey, T. Recombinant electron donors and acceptors to and from reaction center particles, and light dependent menaquinone reduction in isolated membranes of *Heliobacterium modesticaldum*, Arizona State University, 2015.

(66) Nitschke, W.; Mattioli, T.; Rutherford, A. W. In *Origin and Evolution of Biological Energy Conservation*; Baltscheffsky, H., Ed.; VCH: New York, 1996; pp 177–203.

(67) Allen, J. F. *FEBS Lett.* **2005**, *579* (5), 963.

(68) McConnell, M. D.; Cowgill, J. B.; Baker, P. L.; Rappaport, F.; Redding, K. E. *Biochemistry* **2011**, *50* (51), 11034.

(69) Hemm, M. R.; Paul, B. J.; Schneider, T. D.; Storz, G.; Rudd, K. E. *Mol. Microbiol.* **2008**, *70* (6), 1487.

(70) Hansson, A.; Amann, K.; Zygadlo, A.; Meurer, J.; Scheller, H. V; Jensen, P. E. *FEBS J.* **2007**, *274* (7), 1734.

(71) Shi, T.; Bibby, T. S.; Jiang, L.; Irwin, A. J.; Falkowski, P. G. *Mol. Biol. Evol.* **2005**, *22* (11), 2179.

(72) Heathcote, P.; Fyfe, P. K.; Jones, M. R. *Trends in Biochemical Sciences*. 2002, pp 79–87.

(73) Olson, J. M. *Biosystems* **1981**, *14* (1), 89.

(74) Nitschke, W.; Rutherford, A. W. *Trends Biochem. Sci.* **1991**, *16* (7), 241.

(75) Rutherford, A. W.; Mattiolo, T. A.; Nitschke, W. In *Origin and evolution of biological energy conversion*; Baltscheffsky, H., Ed.; New York, 1996; pp 177–204.

(76) Olson, J. M.; Blankenship, R. E. In *Photosynthesis Research*; 2004; Vol. 80, pp 373–386.

(77) Neerken, S.; Amesz, J. *Biochim. Biophys. Acta - Bioenerg.* **2001**, *1507* (1–3), 278.

(78) Zouni, A.; Witt, H.-T.; Kern, J.; Fromme, P.; Krauß, N.; Saenger, W.; Orth, P. *Nature* **2001**, *409* (6821), 739.

(79) Deisenhofer, J.; Epp, O.; Miki, K.; Huber, R.; Michel, H. *Nature* **1985**, *318* (6047), 618.

(80) Lince, M. T.; Vermaas, W. *Eur. J. Biochem.* **1998**, *256* (3), 595.

(81) Schelvis, J. P. M.; van Noort, P. I.; Aartsma, T. J.; van Gorkom, H. J. *Biochim. Biophys. Acta - Bioenerg.* **1994**, *1184* (2–3), 242.

(82) Schweitzer, R. H.; Brudvig, G. W. *Biochemistry* **1997**, *36* (38), 11351.

(83) Schweitzer, R. H.; Melkozernov, A. N.; Blankenship, R. E.; Brudvig, G. W. *J. Phys. Chem. B* **1998**, *102* (42), 8320.

(84) Akiyama, M.; Miyashita, H.; Kise, H.; Watanabe, T.; Mimuro, M.; Miyachi, S.; Kobayashi, M. *Photosynth. Res.* **2002**, *74* (2), 97.

(85) Ohashi, S.; Iemura, T.; Okada, N.; Itoh, S.; Furukawa, H.; Okuda, M.; Ohnishi-Kameyama, M.; Ogawa, T.; Miyashita, H.; Watanabe, T.; Itoh, S.; Oh-oka, H.; Inoue, K.; Kobayashi, M. *Photosynth. Res.* **2010**, *104* (2), 305.

(86) Prince, R. C.; Gest, H.; Blankenship, R. E. *BBA - Bioenerg.* **1985**, *810* (3), 377.

(87) Kobayashi, M.; van de Meent, E. J.; Oh-oka, H.; Inoue, K.; Itoh, S.; Amesz, J. In *Research in Photosynthesis*; Murata, N., Ed.; Kluwer Academic Publishers: Dordrecht, The Netherlands, 1992; pp 393–396.

(88) Poluektov, O. G.; Paschenko, S. V; Utschig, L. M.; Lakshmi, K. V; Thurnauer, M. C. *J. Am. Chem. Soc.* **2005**, *127* (34), 11910.

(89) Santabarbara, S.; Heathcote, P.; Evans, M. C. W. *Biochimica et Biophysica Acta - Bioenergetics*. 2005, pp 283–310.

(90) Muhiuddin, I. P.; Heathcote, P.; Carter, S.; Purton, S.; Rigby, S. E. J.; Evans, M. C. W. *FEBS Lett.* **2001**, *503* (1), 56.

(91) Webber, A. N.; Lubitz, W. *Biochimica et Biophysica Acta - Bioenergetics*. 2001, pp 61–79.

(92) Li, Y.; Lucas, M. G.; Konovalova, T.; Abbott, B.; MacMillan, F.; Petrenko, A.; Sivakumar, V.; Wang, R.; Hastings, G.; Gu, F.; Van Tol, J.; Brunel, L. C.; Timkovich, R.; Rappaport, F.; Redding, K. *Biochemistry* **2004**, *43* (39), 12634.

(93) Chen, M.; Telfer, A.; Lin, S.; Pascal, A.; Larkum, A. W. D.; Barber, J.; Blankenship, R. E. *Photochem. Photobiol. Sci.* **2005**, *4* (12), 1060.

(94) Tomo, T.; Okubo, T.; Akimoto, S.; Yokono, M.; Miyashita, H.; Tsuchiya, T.; Noguchi, T.; Mimuro, M. *Proc. Natl. Acad. Sci. U. S. A.* **2007**, *104* (17), 7283.

(95) Pan, J.; Saer, R. G.; Lin, S.; Beatty, J. T.; Woodbury, N. W. *Biochemistry* **2016**, acs. biochem.6b00317.

(96) Kirmaier, C.; Laporte, L.; Schenck, C. C.; Holten, D. *J. Phys. Chem.* **1995**, *99* (21), 8910.

(97) Dibrova, D. V; Cherepanov, D. A.; Galperin, M. Y.; Skulachev, V. P.; Mulkidjanian, A. Y. *Biochim. Biophys. Acta - Bioenerg.* **2013**, *1827* (11–12), 1407.

(98) Furbacher, P. N.; Tae, G. S.; Cramer, W. A. In *Origin and Evolution of Biological Energy Conservation*; Baltscheffsky, H., Ed.; Wiley-VCH: New York, 1996; pp 221–253.

(99) Bak, D. W.; Elliott, S. J. *Curr. Opin. Chem. Biol.* **2014**, *19* (Supplement C), 50.

(100) Okamura, M. . Y.; Paddock, M. . L.; Graige, M. . S.; Feher, G. *Biochim. Biophys. Acta - Bioenerg.* **2000**, *1458* (1), 148.

(101) Romberger, S. P.; Golbeck, J. H. *Photosynth. Res.* **2012**, *111* (3), 285.

(102) Cardona, T.; Sanchez-Baracaldo, P.; Rutherford, A. W.; Larkum, A. **2017**.

(103) Chauvet, A.; Sarrou, J.; Lin, S.; Romberger, S. P.; Golbeck, J. H.; Savikhin, S.; Redding, K. E. *Photosynth. Res.* **2013**, *116* (1), 1.

(104) Warren, P. V.; Golbeck, J. H.; Warden, J. T. *Biochemistry* **1993**, *32* (3), 849.

(105) Rutherford, A. W.; Osyczka, A.; Rappaport, F. *FEBS Letters*. 2012, pp 603–616.

(106) Berman, H. M.; Westbrook, J.; Zukang, F.; Gilliland, G.; Bhat, T. N.; Weissig, H.; Shindyalov, I. N.; Bourne, P. E. *Nucleic Acids Res.* **2000**, *28* (1), 235.

(107) DeLano, W. L. *Schrödinger LLC* **2014**, http://www.pymol.org.

(108) The Uniprot Consortium. *Nucleic Acids Res.* **2017**, *45* (D1), D158.

(109) Pei, J.; Kim, B.-H. H.; Grishin, N. V. *Nucleic Acids Res.* **2008**, *36* (7), 2295.

(110) Krissinel, E.; Henrick, K. In *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*; 2005; Vol. 3695 LNBI, pp 67–78.

(111) Sievers, F.; Wilm, A.; Dineen, D.; Gibson, T. J.; Karplus, K.; Li, W.; Lopez, R.; McWilliam, H.; Remmert, M.; Söding, J.; Thompson, J. D.; Higgins, D. G. *Mol. Syst. Biol.* **2011**, *7* (1).

(112) Goujon, M.; McWilliam, H.; Li, W.; Valentin, F.; Squizzato, S.; Paern, J.; Lopez, R. *Nucleic Acids Res.* **2010**, *38* (SUPPL. 2).

(113) Kumar, S.; Stecher, G.; Tamura, K. *Mol. Biol. Evol.* **2016**, msw054.

(114) Adachi, J.; Waddell, P. J.; Martin, W.; Hasegawa, M. *J. Mol. Evol.* **2000**, *50* (4), 348.

(115) Whelan, S.; Goldman, N. *Mol. Biol. Evol.* **2001**, *18* (5), 691.

(116) Le, S. Q.; Gascuel, O. *Mol. Biol. Evol.* **2008**, *25* (7), 1307.

(117) Dimmic, M. W.; Rest, J. S.; Mindell, D. P.; Goldstein, R. A. *J. Mol. Evol.* **2002**, *55* (1), 65.

Chapter 5

Hydrogen Production by and Electron Microscopy of *Heliobacterium modesticaldum*

Christopher Gisriel[1], Joseph A. Laureanti[1], Shangji Zhang[1], Savanah D. McMahon[1], Anne K. Jones[1], Kevin E. Redding[1]

[1]School of Molecular Sciences, Arizona State University, Tempe, AZ 85287

Abstract

Biohydrogen has great potential as a clean alternative fuel. Like many members of the Clostridiales order, the genome of the photoheterotroph *Heliobacterium modesticaldum* encodes for multiple hydrogenase- and nitrogenase-related enzymes. However, biohydrogen production in heliobacteria has not previously been observed or characterized. Here, we measured the rate of hydrogen production from *H. modesticaldum* under a variety of conditions and examine its cellular architecture using electron microscopy of both negatively-stained whole cells and of thin-sectioned cells. We conclude that optimal hydrogen production is reached when growth media lack ammonium and is shifted from N_2 to Ar, is grown photoheterotrophically in the light, and has pyruvate in the media. We suggest that *H. modesticaldum* produces hydrogen via photofermentation using its nitrogenase enzyme as the hydrogen production catalyst, making it the first example of this mechanism in a Gram-positive organism, and propose a scheme as the major hydrogen-producing pathway in heliobacteria. In addition, electron microscopy confirms that no membranes form in heliobacteria to accommodate increased levels of the photosynthetic apparatus.

Introduction

The availability of energy sources that humans currently harvest are not only dwindling but their use may also release harmful biproducts which result in climate change, pollution, and health risks [1]. Hydrogen is often considered a desirable alternative fuel because of its lack of harmful byproducts and its high energy density [2]. Currently, most hydrogen is produced by reformation of fossil fuels. Biohydrogen production, the

production of hydrogen by microorganisms, is an appealing alternative, however, because of the versatility of hydrogen synthesis mechanisms among microorganisms. These mechanisms include biophotolysis, dark-fermentation, and photofermentation [1].

The metabolism of microalgae and prokaryotes that perform oxidative photosynthesis can perform biophotolysis of water wherein hydrogen gas is produced rather than carbon-containing biomass. [3] Electrons flow from the two enzymes involved in light-driven electron transfer, Photosystem I (PSI) and Photosystem II (PSII), to those enzymes that are capable of producing hydrogen – hydrogenase and nitrogenase. This is problematic, however, as many hydrogenases are irreversibly inactivated by oxygen, a product of PSII water splitting [4–7]. The use of anaerobic organisms whose photosystems do not produce oxygen is an appealing potential solution to this problem.

Dark-fermentative hydrogen production can be achieved by the conversion of organic compounds, such as sugars, starch, and cellulose, to CO_2 and hydrogen. It has advantages in its high hydrogen evolution rate and non-reliance upon light [8–11]. Excess electrons from the degradation of substrates are routed to hydrogen-producing enzymes to dispose of as a waste product, hydrogen. Photofermentation, however, requires both light and organic compounds. The photosynthetic apparatus of anaerobic bacteria shuttles electrons to nitrogenase where hydrogen is produced in nitrogen-limiting conditions [12].

As alluded to above, two classes of enzymes are typically exploited for their capability to produce hydrogen in biological systems – hydrogenases and nitrogenases. Hydrogenases are ubiquitous throughout all kingdoms of life [13–16]. They are metalloenzymes that catalyze the production of hydrogen, a vital aspect of energy

metabolism [17]. Hydrogenases are divided into two phylogenetically distinct classes based on the composition of their redox-active metal centers: [NiFe], and [FeFe]. They perform the reversible reaction, $H_2 \leftrightarrows 2\ H^+ + 2e^-$, serving as a terminal electron sink for low-potential reductants (proton reduction) or to produce protons and low-potential reductants in energy-yielding processes (hydrogen oxidation) [15].

Nitrogenases perform N_2 fixation, a process that is vital to the biogeochemical nitrogen cycle [18–20]. Nitrogenases are also classified by an active-site metal center involved in their redox chemistry: Mo-, V-, and Fe-forms [21] (molybdenum, vanadium, and iron, respectively). Of these, Mo-nitrogenases have been most highly characterized. Nitrogenase catalyzes the formation of NH_3 from N_2 in an energetically expensive multi-step reaction that can be summarized in the overall reaction (for Mo-nitrogenases), $N_2 + 8\ Fd_{red} + 16\ ATP + 8\ H^+ \rightarrow 2\ NH_3 + H_2 + Fd_{ox} + 16\ ADP + 16\ P_i$ (where Fd_{red} and Fd_{ox} are reduced and oxidized ferredoxin, respectively) [22]. In N_2-deplete conditions, however, active site reduction leads to proton reduction in the overall reaction $4\ ATP + 2\ H^+ + 2\ e^- \rightarrow H_2 + 4\ ADP + 4\ P_i$, essentially causing the nitrogenase to work as a hydrogenase [22,23].

Heliobacteria are the only photosynthetic members of the phylum Firmicutes [24] that are known to use organic carbon sources for photoheterotrophic and chemotrophic growth, producing photosynthetic pigments during both conditions [25]. Electrons are acquired via the oxidation of exogenous carbon compounds, primarily organic acids such as pyruvate. The membrane-anchored cytochrome c_{553} is oxidized by the RC and the electron is transferred into the cell, reducing ferredoxin to drive downstream metabolic

processes. One such outcome may be electron flow towards nitrogenase or hydrogenase(s).

The heliobacterial genome contains genes for the two core Mo-dinitrogenase subunits (*nifD* and *nifK*), dinitrogenase reductase ("Fe protein"; *nifH*), two uptake [NiFe]-hydrogenase-related genes (*hupS*, *hupL*), one [Fe]-hydrogenase-related gene (*hymD*), and potential [FeFe]-hydrogenase-related genes (*nuoEFG*) [25]. Uptake hydrogenases (favoring the unidirectional reaction, $H_2 + A_{ox} \rightarrow 2\ H^+ + A_{red}$) oxidize hydrogen, resulting in NAD^+ reduction to maintain a reduced quinone-pool (Q-pool) via the NADH:MQ oxidoreductase complex (NMOR), therefore the expression of [NiFe]-hydrogenases are less likely to produce H_2. Conversely, proton reduction ($2\ H^+ + A_{red} \rightarrow H_2 + A_{ox}$) by hydrogenase or nitrogenase results in the production of hydrogen gas. Hydrogen metabolism in heliobacteria has not previously been characterized. Heliobacteria are an interesting candidate for the study of hydrogen production, given their light-dependent growth conditions, lack of oxygen-producing enzymes, and genes encoding for hydrogenase and nitrogenase-related activity.

Based on the genome of *H. modesticaldum*, CO_2-fixation pathways have been identified that employ pyruvate:ferredoxin oxidoreductase (PFOR) and phosphoenolpyruvate carboxykinase (PEPCK) [25]. The absence of a full reverse tricarboxylic acid pathway, due to the lack of a functional ATP citrate lyase, likely prevent heliobacteria to perform photoautotrophic growth [25]. Heliobacteria have been shown to fix nitrogen during both phototrophic and chemotrophic growth [25,26]. Nitrogenase activity was detected by acetylene reduction in species that were grown

photoheterotrophically in the presence of a N_2 atmosphere. Some species of heliobacteria have been shown to exhibit a nitrogenase "switch-off" by the addition of ammonium (NH_4^+), similar to that observed in purple and green bacteria [27]. Heliobacteria are the first Gram-positive diazotroph to exhibit a nitrogenase switch-off mechanism [26].

Here, we quantify the presence and rate of hydrogen production in *H. modesticaldum*, in a variety of conditions. We show that *H. modesticaldum* hydrogen production is primarily driven by nitrogenase rather than hydrogenase by making use of its nitrogenase switch-off mechanism. We also show that available pyruvate in the media, heliobacteria's preferred organic carbon source [25], increases the rate and quantity of light-dependent hydrogen production as a sink for excess electrons, thus decreasing the concentration of RCs under these conditions. In pyruvate-deplete conditions, we investigate the possibility that heliobacteria increase their membrane surface area to accommodate the up-regulation of pigment-rich RC complexes by negatively-stained whole cell and thin section transmission electron microscopy (TEM) and show that this is not the case. These TEM images are the first of *H. modesticaldum*, displaying their physiological characteristics in a variety of conditions.

Results and Discussion

Origin of hydrogen production in heliobacteria. The nitrogenase switch-off mechanism occurs when a product of the nitrogenase mechanism, ammonium (NH_4^+), is readily available, and the cell down-regulates nitrogenase expression [26,28]. Our initial hypothesis was that by exploiting the nitrogenase switch-off mechanism, *H. modesticaldum* would maximize its hydrogen production under N_2 limited conditions.

Therefore, if hydrogen is produced by nitrogenase, hydrogen production should be inhibited when heliobacteria is grown in media containing NH_4^+. If, however, hydrogen is produced by hydrogenase, the nitrogenase switch-off should have no effect on hydrogen production unless hydrogenase transcription is also regulated by ammonia.

In the absence of the N_2 substrate, nitrogenase reduces protons to hydrogen ($2\ H^+ + 2\ e^- \rightarrow H_2$), which is coupled to the dephosphorylation of 4 ATP to 4 ADP. Thus, upon the purging of N_2 from the culture, all active nitrogenases should produce H_2 rather than the usual nitrogen fixation to NH_3, significantly increasing the amount of hydrogen produced.

With these hypotheses in mind, we measured hydrogen production in samples that had been purged with N_2 and with Ar, and in the presence or ansence of NH_4^+ (**Scheme 5.1A**). Only samples without NH_4^+ produced hydrogen (**Figure 5.1**). They did so at a maximal average rate of ~0.28 and 1.4 μmol H_2 h^{-1} (mg of cells)$^{-1}$ for N_2 and Ar atmospheres, respectively. Of those, the culture purged with Ar produced the largest amount of hydrogen.

The fact that samples containing NH_4^+ did not produce hydrogen is consistent with the hypothesis that the hydrogen produced by *H. modesticaldum* is solely a result of nitrogenase activity. Had hydrogen been detected in samples with down-regulated nitrogenase transcription (nitrogenase switch-off) it would have been deduced that hydrogenase was actively producing hydrogen. Furthermore, switching from N2 to Ar in the headspace drastically impacted hydrogen production. Because N2 is a substrate for nitrogenase, this also supports the hypothesis that hydrogen production is a result of

nitrogenase activity alone. Alternative hypotheses, however, are that (1) uptake hydrogenase transcription may be upregulated in those experimental conditions where hydrogen was not detected, or (2) that nitrogenase and hydrogenase transcriptional regulation is coupled and dependent upon nitrogenase activity. The presence of low amounts of hydrogen throughout the samples collected in the -NH_4^+, N_2 condition implies that the concentration of uptake hydrogenases, if any, is insufficient to fully oxidize the hydrogen produced. We conclude, then, that the Mo-nitrogenases are the major source of hydrogen production in *H. modesticaldum*.

Mechanism of hydrogen production in heliobacteria. To further investigate which hydrogen production mechanisms employed by heliobacteria, such as biophotolysis or fermentative pathways, we characterized the dependence of hydrogen production in heliobacteria on (1) the energy metabolism mode (i.e. photoheterotrophic or chemotrophic), and (2) the availability of pyruvate as an organic carbon source (**Scheme 5.1B**). Hydrogen production of *H. modesticaldum* without NH_4^+ exhibit strict light-dependence and only produce hydrogen in the presence of light (**Figures 5.2** and **5.3**). The relationship of pyruvate concentration and hydrogen production in the light is approximately linear, with cultures containing 9.1 mM, 18.2 mM, and 72.8 mM pyruvate producing 0.17, 0.40, and 1.9 mmol of hydrogen, respectively (**Figure 5.2**). Cultures containing 9.1 mM, 18.2 mM, and 72.8 mM pyruvate exhibit maximal average rates of 0.62, 1.0, and 1.9 μmol H_2 h^{-1} (mg of cells)$^{-1}$, respectively (**Figure 5.3**).

Only trace amounts of hydrogen were detected from the headspace of cultures grown chemotrophically in the dark, ruling out the possibility of heliobacteria producing

hydrogen by dark-fermentation. Biophotolysis occurs only in oxygenic photosynthesis, which heliobacteria do not perform, ruling this out as well. Coupled with the previously stated observation that hydrogen is produced by nitrogenase, the data presented here are consistent with the hypothesis that *H. modesticaldum* produces hydrogen via photofermentation during photoheterotrophic growth, with the electrons imported in the light reactions of photosynthesis flowing to nitrogenase from ferredoxin. The strict dependence upon the availability of an organic carbon source is also a characteristic of photofermentation, supporting this hypothesis as well.

With the knowledge that electrons are shuttled to nitrogenase in a light-dependent mechanism from pyruvate, we propose a scheme for the major pathway of electrons resulting in the reduction of protons to hydrogen at nitrogenase in heliobacteria in **Figure 5.4**. After uptake via a transmembrane transporter, pyruvate is oxidized by PFOR to acetyl-CoA and CO_2 with the 2 electrons transferred to (two) Fd. Then FNR can catalyze ET from the 2 reduced Fd to NAD(P)+ to produce NAD(P)H, which can be oxidized by NMOR, passing the electrons to the MQ-pool. Electrons can then be shuttled through the membrane to the membrane-anchored cyt c_{553} via the cyt b_6c complex, which oxidizes MQH_2 to MQ, and pumps protons from the N- to the P-side via the Q-cycle. NMOR also pumps protons as MQ is reduced, thus linking the electron transport pathway to ATP production. The RC complex then closes the loop, catalyzing light-dependent ET from cyt c_{553} to ferredoxin. Electrons can be withdrawn from this cycle at the level of Fd_{red} and passed to nitrogenase, where proton reduction results in hydrogen production (red line in **Figure 5.4**). For every electron that is withdrawn from the cycle (e.g. as H_2), an electron

must be supplied (*e.g.* from pyruvate oxidation), explaining the dependence of H_2 production on pyruvate supply.

This is the first example of a Gram-positive organism performing photofermentation, which is perhaps unsurprising because they are also the only photosynthetic member of the phylum Firmicutes [24]. Photofermenting Gram-negative purple non-sulfur bacteria convert light energy into chemical energy resulting in electrons being carried from the RC to the cytochrome *bc* complex via the membrane-contained Q-pool. The electrons can then be transferred to ferredoxin that can reduce nitrogenase leading to hydrogen production in nitrogen-limited conditions. In heliobacteria, however, ferredoxin is directly reduced by the RC, and delivers them to nitrogenase. This removes the cytochrome *bc* complex intermediate, and creates a more direct pathway for photofermentation.

It has been hypothesized that an ancestral heliobacterium received the genes required for photosynthesis via a lateral gene transfer [29]. This ancestral heliobacterium probably already had the machinery to perform dark-fermentation (like all other Firmicutes); ferredoxins were already actively shuttling electrons to nitrogenase and hydrogenase. Upon gaining the photosynthetic gene cluster and expressing its encoded proteins, electrons from the Q-pool could then be coupled to photosynthesis. Whereas this dark-fermentation resulted in the wasteful disposal of carbon as CO_2, this new photosynthetic capability allowed the carbon from pyruvate to be used for biomass.

The ability of *H. modesticaldum* to perform light-dependent hydrogen production was exploited in bio-electrosynthesis experiments [30]. Bio-electrosynthesis is the process

by which microorganisms use electrons from a cathode to form a desired product [31]. In addition to their metabolic characteristics, heliobacteria are ideal candidates for such a study because of their Gram-positive membrane architecture. Because the RCs of heliobacteria are located within their single membrane, and presumably positioned to direct electrons into the cell, the application of cells to an electrode results in close proximity of the electron source, the cathode, and the enzyme that can deliver electrons into the organism's metabolic pathway, the RC. Only a chemical mediator with the ability to penetrate the peptidoglycan cell wall is required to deliver electrons to the RC from the cathode. Laureanti showed that indeed photocurrent is generated in a light-dependent manner, and hydrogen is produced, implying that electrons from the cathode successfully reduce the RC to drive the heliobacterial metabolism [30]. This is the first example of bio-electrosynthesis from a photosynthetic organism.

Qualitative analysis of bacteriochlorophyll g per cell variation. During the process of obtaining hydrogen measurements, each data point included a measurement of the steady-state absorption from 400 – 900 nm (data not shown). The ratio of heliobacteria's major pigment, bacteriochlorophyll *g* (BChl *g*), which exhibits strong absorption at its Q_y band of ~788 nm, is typically be expected to maintain a relatively consistent ratio to the absorbance of a non-pigment-absorbing wavelength like 625 nm that represents only light scattering from cells. This would imply a consistent ratio of RC complexes to cells, as peripheral antenna complexes have not been identified [32,33]. This expectation, however, observed (**Figure 5.5**). Rather, the cultures grown photoheterotrophically in the light appeared to show an increased ratio of BChl *g* content

to cell material as pyruvate became limiting. In the sample grown in the light containing the least pyruvate, 9.1 mM, the ratio of BChl *g* to cells remained consistently high. In the sample grown in the light containing an intermediate value (18.2 mM pyruvate), the ratio began lower and increased around when the culture decreased its hydrogen production, ~46 hours after inoculation. In the sample containing 72.8 mM pyruvate, the ratio remained low. Surprisingly, dark-grown cells with 18.2 mM pyruvate appeared to increase this ratio at a time similar to that of light-grown cells. Therefore, we tentatively conclude that RC concentration is up-regulated upon pyruvate depletion in *H. modesticaldum*. A clear explanation for this phenomenon is lacking, but it is clear that their metabolism suffers as a result of limiting their carbon and electron source, pyruvate.

Many other photosynthetic organisms employ invaginated membranes or entire organelles to accommodate higher amounts of transmembrane antenna or RC complexes for light harvesting, a characteristic that heliobacteria are not known to exhibit. This led to our hypothesis that the membrane surface area of heliobacterial cells increases under pyruvate-limited conditions and was investigated using TEM below.

TEM of H. modesticaldum. An external view of fixed whole cells was acquired by negatively staining *H. modesticaldum* cells (see Materials and Methods). No invaginated membrane features were identified when cells were grown with either 9.1 or 72.8 mM pyruvate (**Figure 5.5**, top left and top right). Neither was there a visible difference between cells fixed at either 24 or 48 hours after inoculation (**Figure 5.5**, bottom left and bottom right). Cell division was commonly identified (**Figure 5.5**, top left and bottom right), and a regularly repeating pattern on cells was observed at higher magnification

settings (**Figure 5.5**, bottom right), likely the matrix of peptidoglycan that constitutes their cell wall. Cells were 5 to 20 µm in length, and ~0.5 µm in width.

To gain an internal view of cellular architecture, cells were fixed, suspended in an epoxy resin, and thin sections were cut with a microtome (see Materials and Methods). Samples were imaged and displayed no evidence of internal membranes or invaginations (**Figure 5.6**). However, cells that were fixed at 48 hours displayed electron-dense spots (**Figure 5.6**, bottom left and bottom right), the identity of which remains unknown, but do not appear to be surrounded by a lipid bilayer. Only 1-2 dark spots were identified in many cells, making it unlikely that these spots are an antenna-rich structural feature that is upregulated for light-harvesting when carbon sources become limiting. Perhaps more convincingly, these dark spots hold no resemblance to EM images of chlorosomes from green sulfur bacteria [34] or invaginated membranes of purple bacteria [35].

In conclusion, *H. modesticaldum* do not appear to increase their membrane surface area to accommodate more RCs. The RC concentration in many photosynthetic organisms is regulated dependent upon the availability of light that can be converted to useful chemical energy. Therefore, it is surprising that the data presented in **Figure 5.5** suggest that regulation of RC concentration is not as important; cultures grown in light and dark conditions exhibit similar ratios of BChl *g* to cell density. This implies that photosynthetic activity is not heavily relied upon for metabolism. Regardless, it appears that photoheterotrophic growth is preferred over chemotrophic growth, although chlorophyll-like features remain in the absorption spectrum of whole cells even when grown chemotrophically in the dark [25]. This versatility probably reflects their ecological

niche, thriving in volcanic soils at high temperatures, where light and carbon sources probably vary in the environment dependent upon the metabolic requirements of other organisms in which they share symbiosis. Although it is tempting to imagine heliobacterial metabolism switching distinctively from photoheterotrophic to chemotrophic modes, their environment probably allows very little light to reach their RCs. Consequently, a more realistic view of heliobacterial metabolism may be a constant balance between photoheterotrophic and chemotrophic growth, with only slight metabolic shifts reliant upon the availability of light during the diurnal cycle.

Materials and Methods

Growth. *H. modesticaldum* was grown anaerobically in variations of pyruvate yeast extract (PYE) as described previously [28]. Light-grown cultures were given ~33 µE of 810 nm light at 50 °C. Where pyruvate was varied in the media, either 9.1 mM, 18.2 mM or 72.8 mM was used, where 18.2 mM is the optimized literature value [26]. Where NH_4^+ was varied, NH_4^+ was either not added to the media or in the amount of 7.6 mM.

Biohydrogen measurements. Hydrogen production was measured by gas chromatography (GC). Headspace was sampled using a gas-tight syringe purged with Ar. 250 µL of the culture headspace was removed and injected into the GC for separation by a thermal conductivity detector (TCD) in Ar carrier gas in an SRI model 310 with a 6' molecular sieve 13X packed column. The GC-TCD was calibrated by calculating the peak areas of a set of known hydrogen concentrations to which a standard curve was fit for the calculation of actual hydrogen concentration in culture headspace. The ideal gas law was used to calculate the molar concentration of hydrogen.

Negative Stain TEM. Cells were fixed by incubation in 25% formaldehyde for 30 minutes. Cells were washed and resuspended in 50 mM MOPS buffer, pH=7. Cells were applied to 300-mesh carbon-formvar coated copper grids for 1 minute. They were stained with a 2% uranyl acetate solution for 1 minute. Images were generated with a Philips CM12 TEM operated at 80 kV. Images were acquired with a Gatan model 791 CCD camera (1024 x 1024 array) cooled to -30 °C.

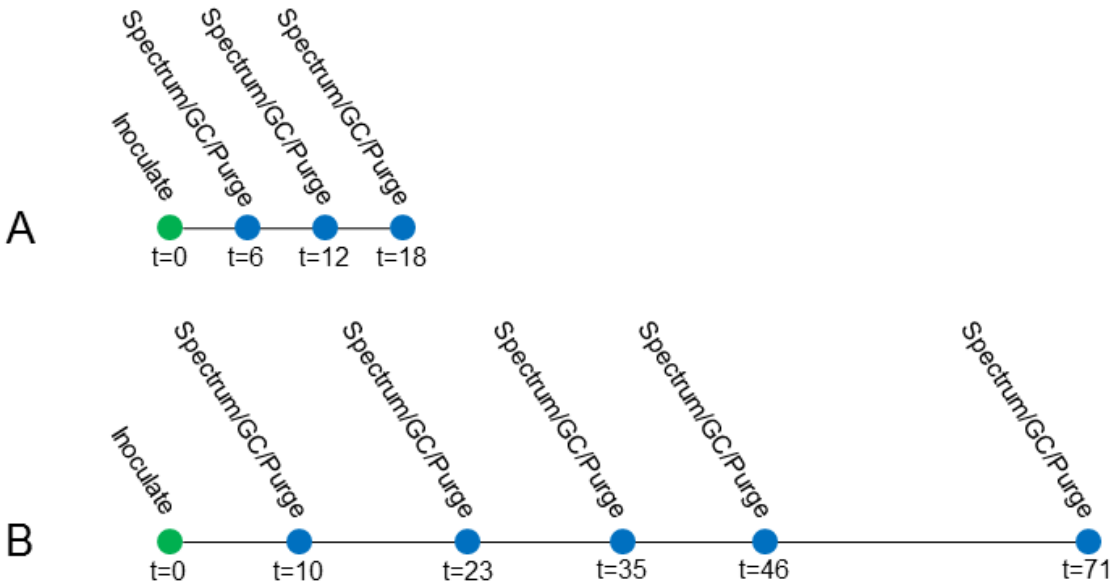
Thin section TEM. Cells were fixed by incubation in 20% glutaraldehyde for 1 hour. Cells were washed 3 times and resuspended in 20 mM sodium phosphate buffer, pH=7. Cells were pelleted and agarose was added to 1%. Solidified agarose pellet balls were washed 5 times with 20 mM sodium phosphate buffer, pH=7.0. Pellet balls were fixed with 1% osmium tetroxide for 1.5 hours. Pellet balls were washed three times with 20 mM sodium phosphate buffer, pH=7, and three times with water. Pellet balls were *en bloc* stained with 0.5% uranyl acetate overnight at 4 °C and washed with water five tims. Pellet balls were dehydrated through a graded series of acetone solutions (20%, 40%, 60%, 80%) for 30 minutes in each incubation step. Then, dehydrated pellet balls were incubated in 100% acetone three times for 1.5 hours. They were infiltrated with 25% epoxy resin for 1 hour, 50% epoxy resin for 3 hours, and 75% epoxy resin overnight. Then, pellet balls were changed into 100% epoxy resin six times, each for 12 hours. Fully infiltrated agarose pellet balls were embedded with 100% epoxy resin in a 60 °C oven until resin blocks were polymerized. Solidified resin blocks were trimmed under a dissecting microscope with a double-edged razor blade. The sample surface area was exposed on a trapezoidal surface in a pyramid body block shape. Once the sample surface

area was limited to a base edge of the trapezoidal surface that was smaller than 500 μm, the sample block was further trimmed with 100 nm thin sections to polish the sample surface via an Ultracut-E ultramicrotome (Reichert-Jung, the US). Then, 70 nm thin serial section samples were trimmed by an Ultracut-R ultramicrotome (Leica Microsystems, Austria). The thin serial sections were picked up via Formvar-coated EM cooper slot grids (2 mm x 1 mm). Sample grids were double contrast stained with 0.5% uranyl acetate and 3% lead citrate. The grids were imaged with a JEOL 1200EX TEM transmission electron microscope (JEOL, Japan) operating at 80 KV.

Acknowledgements

This research was supported by the Arizona State University Lightworks grant awarded to K. E. Redding, A. K. Jones, and S. C. Popat.

Schemes



Scheme 5.1 Time sequence of hydrogen production experiments. **A** scheme was used in the nitrogenase shutoff experiments in which NH_4^+ and atmosphere were varied. **B** scheme was used in the hydrogen optimization experiments (**Figures 5.2** and **5.3**). Media was inoculated (green circle) with a 1/100 dilution of cells from a culture in late log phase. At the time points indicated (blue circles, labelled in units of hours), the spectrum of the cells was collected, the hydrogen in the headspace was measured (**Figures 5.1 and 5.2**), and the headspace was purged with the initial atmosphere gas for 5 minutes.

Figures

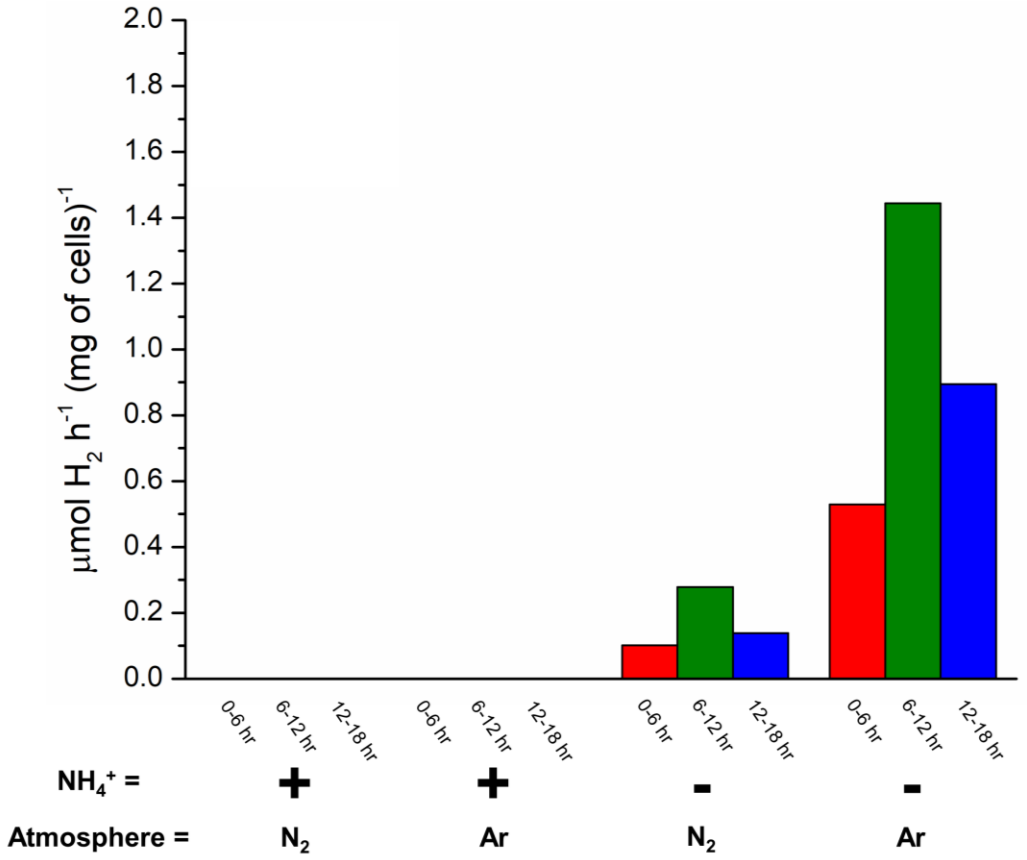


Figure 5.1 Average rate of hydrogen production from heliobacterial cultures varying NH_4^+ and atmosphere. Media either contained (bold "+") or did not contain (bold "-") NH_4^+. In addition, media was purged with either N_2 or Ar gas before initiating H_2 production and after each H_2 measurement at 6, 12, and 18 hours after the experiment had started (see **Scheme 5.1**). Hydrogen production is plotted as an average rate of hydrogen production (µmol H_2 h^{-1} (mg of cells)$^{-1}$) at each time point since the previous measurement.

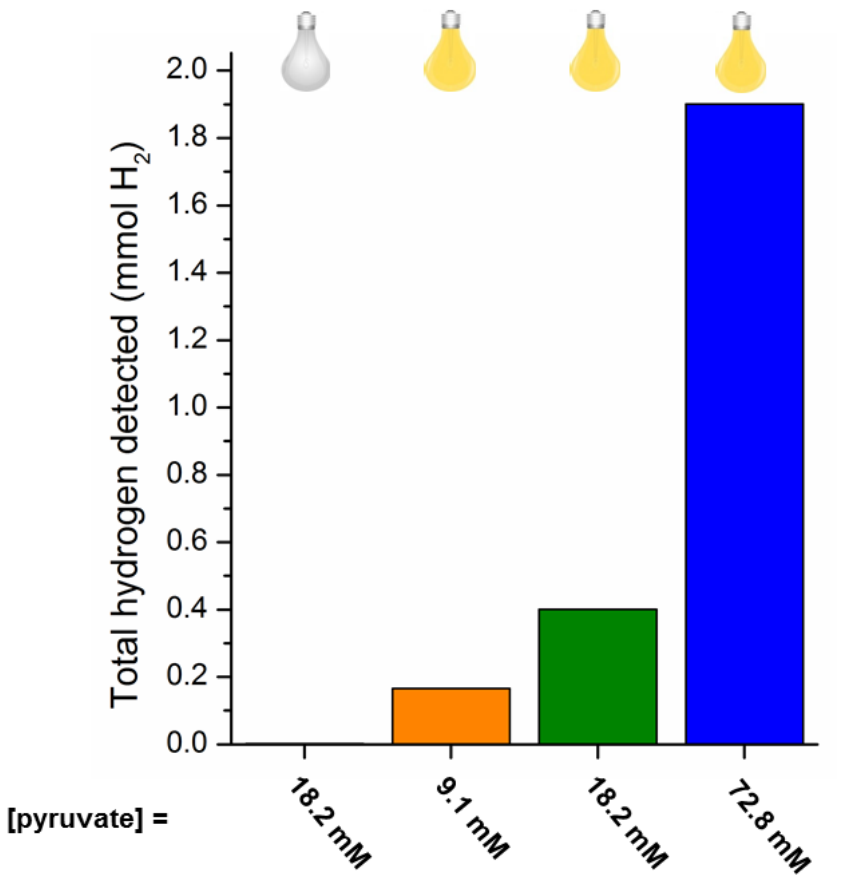


Figure 5.2 Total hydrogen produced by heliobacterial cultures. Heliobacterial cultures were grown in varying amounts of pyruvate (bottom, bold labels) and with or without an 810 nm light source (signified by the top light bulbs). Total hydrogen detected from each culture since inoculation is displayed.

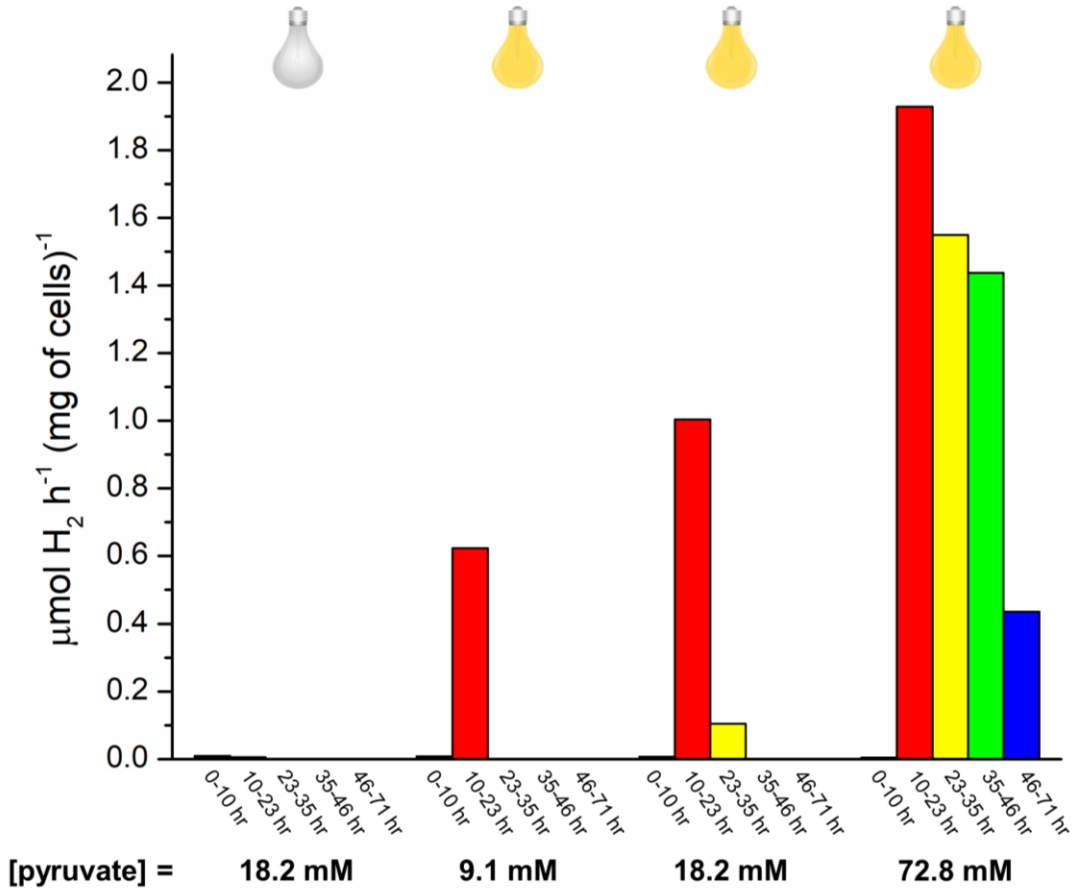


Figure 5.3 Average rates of hydrogen production in heliobacterial cultures. Heliobacterial cultures were grown in varying amounts of pyruvate (bottom, bold labels) and with or without an 810 nm light source (signified by the top light bulbs). Hydrogen production was measured at various time points and plotted as an average rate of hydrogen production (µmol H_2 h^{-1} (mg of cells)$^{-1}$) since the previously measured time point.

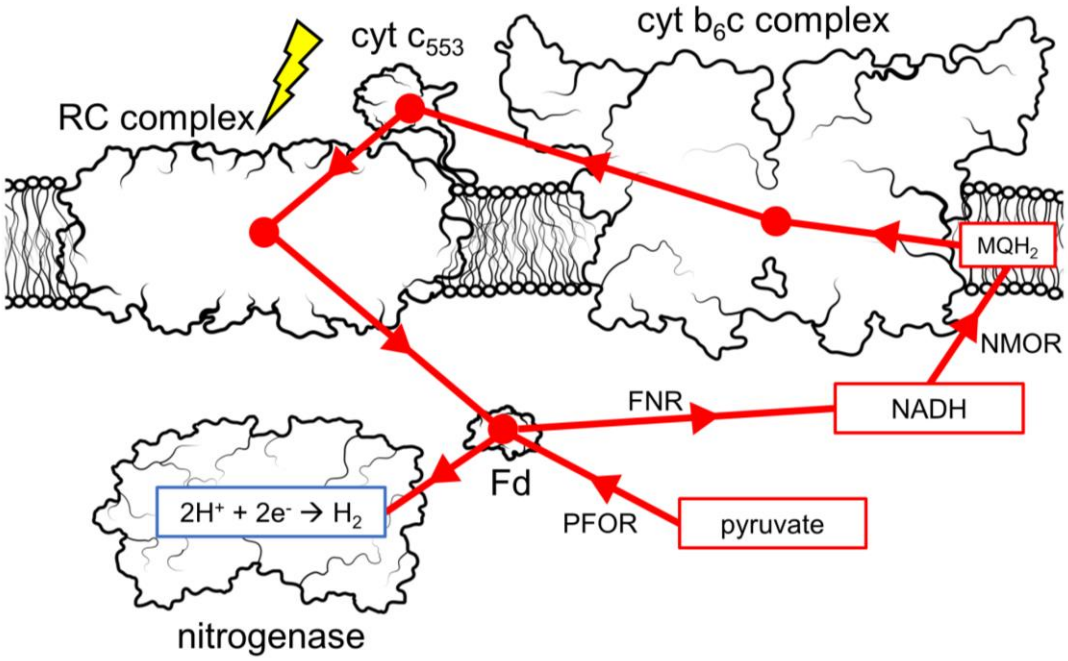


Figure 5.4 Major path of electron transport resulting in hydrogen production in heliobacteria. Electrons (red arrows) from pyruvate reduce ferredoxin (PFOR) which can either directly reduce nitrogenase or electrons can reduce NAD+ to NADH via ferredoxin:NAD^+ oxidoreductase (FNR). NADH reduced MQ via the NADH:MQ oxidoreductase (NMOR) and electrons travel through the cyt b_6c complex, cyt c_{553}, RC complex, ferredoxin, and to nitrogenase to reduce protons to hydrogen.

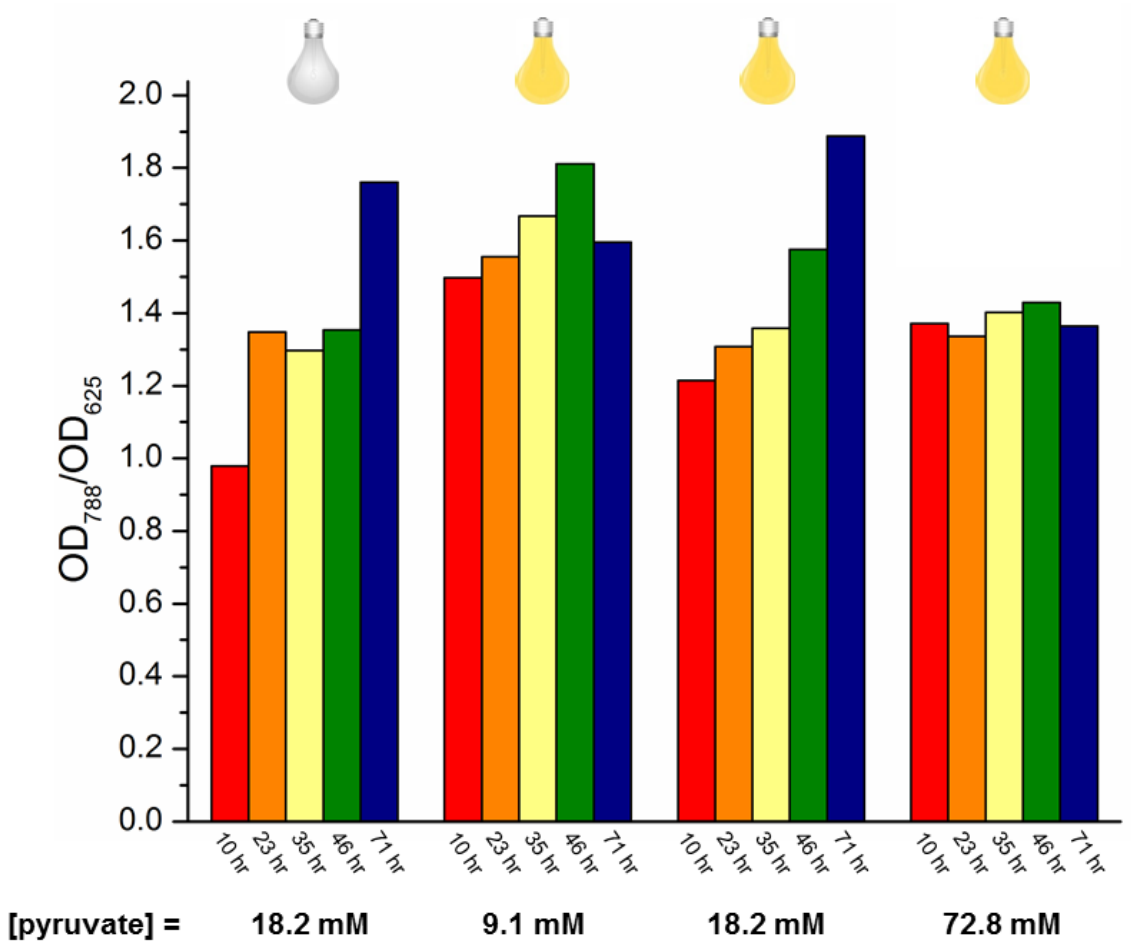


Figure 5.5 Ratio of BChl *g* content to cell density in heliobacteria cultures. Cells were grown in the dark (dark bulb) or light (bright bulb) in medium with the indicated initial concentrations of pyruvate. The absorption at the Q_y band of BChl *g* (788 nm) and a non-absorbing wavelength (625 nm) were collected. Differences in the ratio of the two imply that the number of RCs per cell varies.

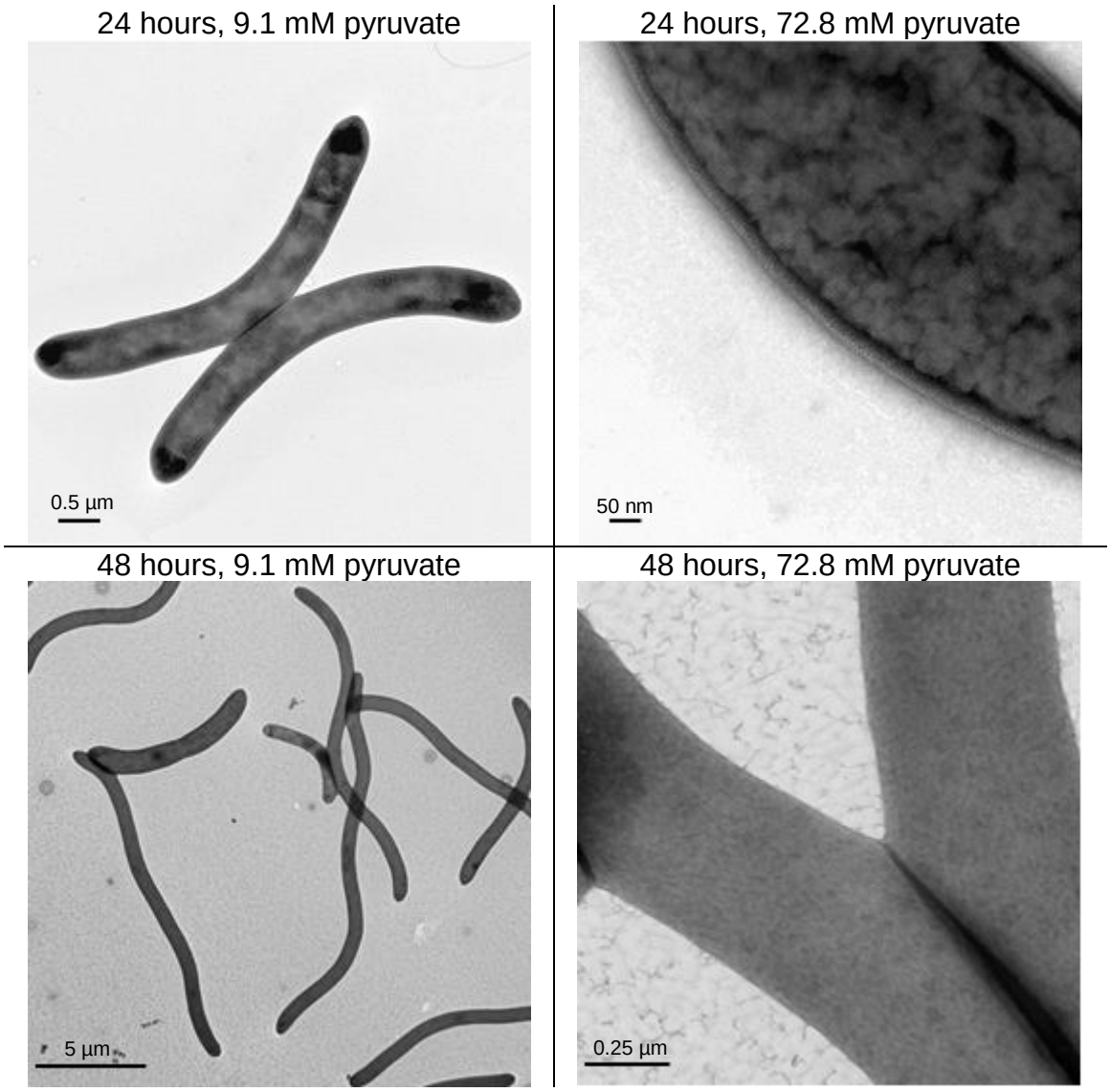


Figure 5.6 Negatively-stained *H. modesticaldum* cells grown without NH_4^+. Panels are titled with the time point at which the cells were fixed and the initial concentration of pyruvate in which they were grown.

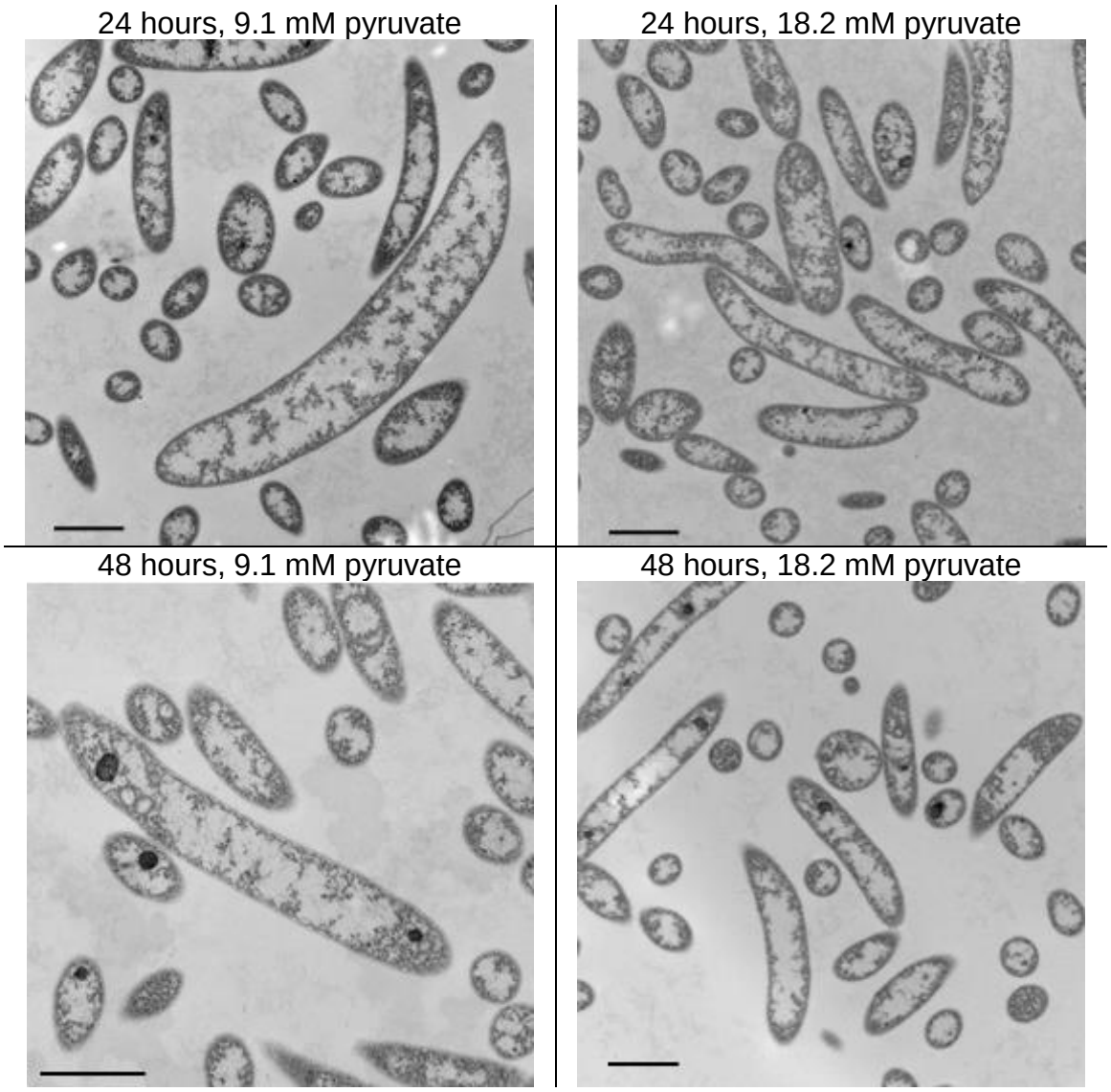


Figure 5.7 Thin-sectioned *H. modesticaldum* cells grown without NH_4^+. Panels are titled with the time point at which the cells were fixed and the initial concentration of pyruvate in which they were grown. All scale bars are 1.0 μm.

References

(1) Das, D.; Veziroglu, T. N. *Int. J. Hydrogen Energy* **2001**, *26* (1), 13.

(2) Levin, D. B.; Pitt, L.; Love, M. *Int. J. Hydrogen Energy* **2004**, *29* (2), 173.

(3) Benemann, J. R. *Int. J. Hydrogen Energy* **1997**, *22* (10–11), 979.

(4) Sarkar, S.; Pandey, K. D.; Kashyap, A. K. *Int. J. Hydrogen Energy* **1992**, *17* (9), 689.

(5) Lambert, G. R.; Smith, G. D. *Arch. Biochem. Biophys.* **1980**, *205* (1), 36.

(6) Tel-Or, E.; Luijk, L. W.; Packer, L. *FEBS Lett.* **1977**, *78* (1), 49.

(7) Bothe, H.; Distler, E.; Eisbrenner, G. *Biochimie* **1978**, *60* (3), 277.

(8) Kumar, N.; Das, D. *Process Biochem.* **2000**, *35* (6), 589.

(9) Tanisho, S.; Wakao, N.; Kosako, Y. *J. Chem. Eng. Jpn.* **1983**, *16* (6), 529.

(10) Tanisho, S.; Suzuki, Y.; Wakao, N. *Int. J. Hydrogen Energy* **1987**, *12* (9), 623.

(11) Brosseau, J. D.; Zajic, J. E. *Int. J. Hydrogen Energy* **1982**, *7* (8), 623.

(12) Luque, R.; Lin, C. S.; Wilson, K.; Clark, J. Elsevier Inc., 2010; pp 313–314.

(13) Vignais, P. M.; Billoud, B.; Meyer, J. *FEMS Microbiology Reviews*. 2001, pp 455–501.

(14) Meyer, J. *Cellular and Molecular Life Sciences*. 2007, pp 1063–1084.

(15) Vignais, P. M.; Billoud, B. *Chem. Rev.* **2007**, *107* (10), 4206.

(16) Posewitz, M. C.; Mulder, D. W.; Peters, J. W. *Curr. Chem. Biol.* **2008**, *2* (2), 178.

(17) Mulder, D. W.; Shepard, E. M.; Meuser, J. E.; Joshi, N.; King, P. W.; Posewitz, M. C.; Broderick, J. B.; Peters, J. W. *Structure*. 2011, pp 1038–1052.

(18) Canfield, D. E.; Glazer, A. N.; Falkowski, P. G. *Science.* **2010**, *330* (6001), 192.

(19) Thamdrup, B. *Annu. Rev. Ecol. Evol. Syst.* **2012**, *43* (1), 407.

(20) Gruber, N.; Galloway, J. N. *Nature* **2008**, *451* (7176), 293.

(21) Peters, J. W.; Szilagyi, R. K. *Current Opinion in Chemical Biology*. 2006, pp 101–108.

(22) Hoffman, B. M.; Lukoyanov, D.; Yang, Z.-Y.; Dean, D. R.; Seefeldt, L. C. *Chem. Rev.* **2014**, *114* (8), 4041.

(23) Burgess, B. K.; Lowe, D. J. *Chem. Rev.* **1996**, *96* (7), 2983.

(24) Bryantseva, I. A.; Gorlenko, V. M.; Kompantseva, E. I.; Achenbach, L. A.; Madigan, M. T. *Arch. Microbiol.* **1999**, *172* (3), 167.

(25) Tang, K.-H.; Yue, H.; Blankenship, R. E. *BMC Microbiol.* **2010**, *10*, 150.

(26) Kimble, L. K.; Madigan, M. T. *Arch. Microbiol.* **1992**, *158*, 155.

(27) Ludden, P. W.; Roberts, G. P.; Lowery, R. G.; Fitzmaurice, W. P.; Saari, L. L.; Lehman, L.; Lies, D.; Woehle, D.; Wirt, H.; Murrell, S. A. *Nitrogen Fixat. hundred years after Proc. 7th Int. Congr. N [Triple-bond] Nitrogen Fixat. Koln (Cologne), F.R.G., March 13-20, 1980 / Ed. by H. Bothe, F.J. Bruijn W.E. Newt.* **1988**.

(28) Kimble, L. K.; Mandelco, L.; Woese, C. R.; Madigan, M. T. *Arch. Microbiol.* **1995**, *163* (4), 259.

(29) Xiong, J.; Inoue, K.; Bauer, C. E. *Proc. Natl. Acad. Sci. U. S. A.* **1998**, *95* (25), 14851.

(30) Laureanti, J. A. Bioinspired Electrocatalytic Hydrogen Production: Synthetic and Biological Approaches, Arizona State University, 2017.

(31) Nevin, K. P.; Hensley, S. A.; Franks, A. E.; Summers, Z. M.; Ou, J.; Woodard, T. L.; Snoeyenbos-West, O. L.; Lovley, D. R. *Appl. Environ. Microbiol.* **2011**, *77* (9), 2882.

(32) Fuller, R. C.; Sprague, S. G.; Gest, H.; Blankenship, R. E. *FEBS Lett.* **1985**, *182* (2), 345.

(33) Miller, K. R.; Jacob, J. S.; Smith, U.; Kolaczkowski, S.; Bowman, M. K. *Arch. Microbiol.* **1986**, *146* (2), 111.

(34) Kudryashev, M.; Aktoudianaki, A.; Dedoglou, D.; Stahlberg, H.; Tsiotis, G. *Biochim. Biophys. Acta - Bioenerg.* **2014**, *1837* (10), 1635.

(35) Okubo, Y.; Futamata, H.; Hiraishi, A. *Appl. Environ. Microbiol.* **2006**, *72* (9), 6225.

Chapter 6

Conclusions

Christopher Gisriel[1]

[1]School of Molecular Sciences, Arizona State University, Tempe AZ 85287

Summary and Conclusions

Improving the purification procedure for the heliobacterial reaction center optimized the homogeneity of the resultant oligomeric state of the pure protein. This purification was used in ultrafast transient absorption spectroscopy experiments to identify the rates of electron transfer events between cofactors of the reaction center and the presence of an accessory bacteriochlorophyll *g* was detected for the first time. The initial charge separation rate exhibited similarity to the electron transfer rates of the heliobacterial reaction center's closest homolog, Photosystem I, but the rate of reduction of the terminal electron acceptor is different, suggesting a different composition and positioning of electron transfer cofactors.

The optimized purification technique also allowed for the successful crystallization of the heliobacterial reaction center. Heliobacterial reaction center crystals were used in X-ray crystallography experiments and a 2.2-Å resolution crystal structure was solved. This crystal structure holds immense significance in the field of structural biology and evolution as they pertain to photosynthesis which is summarized in the following paragraphs.

First, a homodimeric ancestor preceded all known heterodimeric reaction centers. Therefore, the homodimeric heliobacterial reaction center exhibits a characteristic central to that of extinct homodimeric reaction centers, providing new insight into how they may have functioned. The heliobacterial reaction center has no subunits other than the two single transmembrane helix subunits, the newly discovered PshX. The structure reveals that the surface is less charged than that of other reaction centers. These aspects of the

heliobacterial reaction center imply that electron donors and acceptors to homodimeric reaction centers were and are today less based upon charged interactions than that of heterodimeric reaction centers; they are governed more by hydrophobic and (di)polar interactions.

Second, the HbRC represents a 4th class of reaction center to have its structure solved, that is, it is the bacterial Type I counterpart to Photosystem I involved in oxygenic photosynthesis whose structure was solved more than 20 years ago. Consequently, it provides a new set of data by which to compare both Type I and Type II reaction center structures. The heliobacterial reaction center structure shows that the ET domain of all reaction centers is a highly conserved motif. Surprisingly, the heliobacterial reaction center structure reveals that its antenna domain transmembrane helices are just as similar to the analogous domain in Photosystem I as they are to the CP43 and CP47 subunits associated with the Type II reaction center, Photosystem II. This implies that the antenna-coordinating 6-TMH arrangement is conserved from a protein ancestral to all extant antenna domains.

Third, the heliobacterial reaction center is the only known reaction center that does not require a quinone for forward electron transfer, a characteristic that is ubiquitous to all other known reaction centers. Whereas at least 3 cofactors are found between the initial electron donor and terminal electron acceptor in other reaction centers, the heliobacterial reaction center requires only two. These cofactors are spread more evenly across the membrane to accommodate electron transfer. In addition, the hydrophobic thicknesses of the known reaction center structures were calculated and the heliobacterial

reaction center displayed this distance to be ~4 Å shorter than the other reaction centers, another accommodation contributing to successful electron transfer from ultimate donor to acceptor.

Although a quinone does not appear in the structure nor is it essential for forward electron transfer to F_X, we hypothesize that under some circumstances a mobile quinone may be used as an alternate terminal electron acceptor, but that the crystal structure was stripped of this loosely bound quinone during isolation. If the heliobacterial reaction center does have the ability to reduce mobile quinones, the Type I/II nomenclature that is in common use is inadequate. Because the heliobacterial reaction center, purple bacterial reaction center, and Photosystem II all can reduce mobile quinones, we propose that the ancestral reaction center did as well. Furthermore, the ancestor of Type I reaction centers probably did as well. We posit that the lineage of reaction centers that led to modern Photosystem I fixed its quinone and heterodimerized to decrease the probability of producing highly reactive singlet oxygen upon charge recombination from a high-energy cofactor. The heliobacterial reaction center, however, probably remained relatively similar to the ancestor it shares with Photosystem I because oxygen tolerance was never a requirement.

Finally, hydrogen production in *Heliobacterium modesticaldum* was quantified for the first time. It was discovered that *H. modesticaldum* perform photofermentation. This is the first example of photofermentation is any Gram-positive bacteria. Transmission electron microscopy was used to image negatively stained whole cells and thin-sectioned cells, the latter of which had not previously been performed on *H.*

modesticaldum, and neither of which had been performed under a variety of conditions. It was confirmed that under none of the conditions in which the cells were grown did the heliobacteria extend their membrane surface area to accommodate a higher concentration of reaction centers.

Bibliography

CHAPTER 1

(1) Bryant, D. A.; Frigaard, N.-U. U. *Prokaryotic photosynthesis and phototrophy illuminated*; 2006; Vol. 14, pp 488–496.

(2) Schubert, W. D.; Klukas, O.; Saenger, W.; Witt, H. T.; Fromme, P.; Krauss, N. *J. Mol. Biol.* **1998**, *280* (2), 297.

(3) Gest, H. *Photosynth. Res.* **1994**, *41* (1), 17.

(4) Gest, H.; Favinger, J. L. *Arch. Microbiol.* **1983**, *136* (1), 11.

(5) Woese, C. R.; Debrunner-Vossbrinck, B. a; Oyaizu, H.; Stackebrandt, E.; Ludwig, W. *Science* **1985**, *229*, 762.

(6) Beck, H.; Hegeman, G. D.; White, D. *FEMS Microbiol. Lett.* **1990**, *69* (3), 229.

(7) Vermaas, W. F. J. *Photosynth. Res.* **1994**, *41* (1), 285.

(8) Deisenhofer, J.; Epp, O.; Miki, K.; Huber, R.; Michel, H. *Nature* **1985**, *318* (6047), 618.

(9) Krauss, N.; Hinrichs, W.; Witt, I.; Fromme, P.; Pritzkow, W.; Dauter, Z.; Betzel, C.; Wilson, K. S.; Witt, H. T.; Saenger, W. *Nature* **1993**, *361* (6410), 326.

(10) Zouni, A.; Witt, H. T.; Kern, J.; Fromme, P.; Krauss, N.; Saenger, W.; Orth, P. *Nature* **2001**, *409* (6821), 739.

(11) Blankenship, R. E. *Photosynth. Res.* **1992**, *33*, 91.

(12) Blankenship, R. E. *Molecular Mechanisms of Photosynthesis*, 2nd ed.; John Wiley & Sons, Ltd.: Southern Gate, Chicester, West Sussex, UK, 2008.

(13) Nitschke, W.; Rutherford, A. W. *Trends Biochem. Sci.* **1991**, *16* (7), 241.

(14) Fischer, W. W.; Hemp, J.; Johnson, J. E. *Annu. Rev. Earth Planet. Sci.* **2016**, *44* (1), 647.

(15) Gupta, R. S. *Photosynthesis Research*. 2003, pp 173–183.

(16) Bryant, D. A.; Liu, Z.; Li, T.; Zhao, F.; Costas, A. M. G.; Klatt, C. G.; Ward, D. M.; Frigaard, N.-U.; Overmann, J. In *Functional genomics and evolution of photosynthetic systems*; Springer Netherlands, 2012; pp 47–102.

(17) Sousa, F. L.; Shavit-Grievink, L.; Allen, J. F.; Martin, W. F. *Genome Biol. Evol.*

2013, *5* (1), 200.

(18) Brockmann, H.; Lipinski, A.; Woese, C.; Madigan, M. *Arch. Microbiol.* **1983**, *136* (1), 17.

(19) Michalski, T. J.; Hunt, J. E.; Bowman, M. K.; Smith, U.; Bardeen, K.; Gest, H.; Norris, J. R.; Katz, J. J. *Proc. Natl. Acad. Sci. U. S. A.* **1987**, *84* (9), 2570.

(20) Cardona, T. *Photosynthesis Research*. 2015, pp 111–134.

(21) Tang, K.-H.; Yue, H.; Blankenship, R. E. *BMC Microbiol.* **2010**, *10*, 150.

(22) Kimble, L. K.; Madigan, M. T. *Arch. Microbiol.* **1992**, *158*, 155.

CHAPTER 2

(1) Trost, J. T.; Blankenship, R. E. *Biochemistry* **1989**, *28* (26), 9898.

(2) Sarrou, I.; Khan, Z.; Cowgill, J.; Lin, S.; Brune, D.; Romberger, S.; Golbeck, J. H.; Redding, K. E. *Photosynth. Res.* **2012**, *111* (3), 291.

(3) Liebl, U.; Mockensturm-Wilson, M.; Trost, J. T.; Brune, D. C.; Blankenship, R. E.; Vermaas, W. F. J. *Proc. Natl. Acad. Sci. U. S. A.* **1993**, *90* (15), 7124.

(4) Brockmann, H.; Lipinski, A.; Woese, C.; Madigan, M. *Arch. Microbiol.* **1983**, *136* (1), 17.

(5) Kobayashi, M.; Watanabe, T.; Ikegami, I.; van de Meent, E. J.; Amesz, J. *FEBS Lett.* **1991**, *284* (1), 129.

(6) van de Meent, E. J.; Kobayashi, M.; Erkelens, C.; van Veelen, P. A.; Amesz, J.; Watanabe, T. *BBA - Bioenerg.* **1991**, *1058* (3), 356.

(7) Krauss, N.; Hinrichs, W.; Witt, I.; Fromme, P.; Pritzkow, W.; Dauter, Z.; Betzel, C.; Wilson, K. S.; Witt, H. T.; Saenger, W. *Nature* **1993**, *361* (6410), 326.

(8) Chauvet, A.; Sarrou, J.; Lin, S.; Romberger, S. P.; Golbeck, J. H.; Savikhin, S.; Redding, K. E. *Photosynth. Res.* **2013**, *116* (1), 1.

(9) Ferlez, B.; Cowgill, J.; Dong, W.; Gisriel, C.; Lin, S. S.; Flores, M.; Walters, K.; Cetnar, D.; Redding, K. E. K. E.; Golbeck, J. H. J. H. *Biochemistry* **2016**, *55* (16), 2358.

(10) Kleinherenbrink, F. A. M.; Aartsma, T. J.; Amesz, J. *BBA - Bioenerg.* **1991**, *1057* (3), 346.

(11) Heinnickel, M.; Agalarov, R.; Svensen, N.; Krebs, C.; Golbeck, J. H.; Park, U. V; Pennsyl, V. *Biochemistry* **2006**, *45* (21), 6756.

(12) Nuijs, A. M.; van Dorssen, R. J.; Duysens, L. N. M.; Amesz, J. *Proc. Natl. Acad. Sci. U. S. A.* **1985**, *82* (20), 6865.

(13) Takaichi, S.; Inoue, K.; Akaike, M.; Kobayashi, M.; Oh-oka, H.; Madigan, M. T. *Arch. Microbiol.* **1997**, *168* (4), 277.

(14) Kleinherenbrink, F. A. M.; Ikegami, I.; Hiraishi, A.; Otte, S. C. M.; Amesz, J. *Biochim. Biophys. Acta - Bioenerg.* **1993**, *1142* (1–2), 69.

(15) Brok, M.; Vasmel, H.; Horikx, J. T. G.; Hoff, A. J. *FEBS Lett.* **1986**, *194* (2), 322.

(16) Muhiuddin, I. P.; Rigby, S. E. J.; Evans, M. C. W.; Amesz, J.; Heathcote, P. *Biochemistry* **1999**, *38* (22), 7159.

(17) Brettel, K.; Leibl, W.; Liebl, U. *Biochimica et Biophysica Acta - Bioenergetics*. 1998, pp 175–181.

(18) Miyamoto, R.; Mino, H.; Kondo, T.; Itoh, S.; Oh-oka, H. *Biochemistry* **2008**, *47* (15), 4386.

(19) Kondo, T.; Matsuoka, M.; Azai, C.; Mino, H.; Oh-oka, H.; Itoh, S. Springer, Berlin, Heidelberg, 2013; pp 21–24.

(20) McConnell, M. D.; Sun, J.; Siavashi, R.; Webber, A.; Redding, K. E.; Golbeck, J. H.; Van Der Est, A. *Biochim. Biophys. Acta - Bioenerg.* **2015**, *1847* (4–5), 429.

(21) Brettel, K.; Leibl, W. *Biochimica et Biophysica Acta - Bioenergetics*. 2001, pp 100–114.

(22) Lancaster, C. R. D.; Michel, H.; Lancaster, C. R. D.; Michel, H. In *Handbook of Metalloproteins*; John Wiley & Sons, Ltd: Chichester, 2006.

(23) Page, C. C.; Moser, C. C.; Chen, X.; Dutton, P. L. *Nature* **1999**, *402*, 47.

(24) Kleinherenbrink, F. A. M.; Amesz, J. *BBA - Bioenerg.* **1993**, *1143* (1), 77.

(25) Pan, J.; Lin, S.; Allen, J. P.; Williams, J. C.; Frank, H. A.; Woodbury, N. W. *J. Phys. Chem. B* **2011**, *115* (21), 7058.

CHAPTER 3

(1) Schramski, J. R.; Gattie, D. K.; Brown, J. H. *Proc. Natl. Acad. Sci. U. S. A.* **2015**, *112* (31), 9511.

(2) Schubert, W. D.; Klukas, O.; Saenger, W.; Witt, H. T.; Fromme, P.; Krauss, N. *J. Mol. Biol.* **1998**, *280* (2), 297.

(3) Bryant, D. A.; Frigaard, N.-U. U. *Prokaryotic photosynthesis and phototrophy illuminated*; 2006; Vol. 14, pp 488–496.

(4) Gest, H.; Favinger, J. L. *Arch. Microbiol.* **1983**, *136* (1), 11.

(5) Fuller, R. C.; Sprague, S. G.; Gest, H.; Blankenship, R. E. *FEBS Lett.* **1985**, *182* (2), 345.

(6) Trost, J. T.; Blankenship, R. E. *Biochemistry* **1989**, *28* (26), 9898.

(7) Cardona, T. *Photosynthesis Research*. 2015, pp 111–134.

(8) Liebl, U.; Mockensturm-Wilson, M.; Trost, J. T.; Brune, D. C.; Blankenship, R. E.; Vermaas, W. F. J. *Proc. Natl. Acad. Sci. U. S. A.* **1993**, *90* (15), 7124.

(9) Büttner, M.; Xie, D. L.; Nelson, H.; Pinther, W.; Hauska, G.; Nelson, N. *Biochim. Biophys. Acta* **1992**, *1101* (2), 154.

(10) Bryant, D. A.; Costas, A. M. G.; Maresca, J. A.; Chew, A. G. M.; Klatt, C. G.; Bateson, M. M.; Tallon, L. J.; Hostetler, J.; Nelson, W. C.; Heidelberg, J. F.; Ward, D. M. *Science* **2007**, *317* (5837), 523.

(11) Blankenship, R. E. *Photosynth. Res.* **1992**, *33*, 91.

(12) Jordan, P.; Fromme, P.; Witt, H. T.; Klukas, O.; Saenger, W.; Krauss, N. *Nature* **2001**, *411* (6840), 909.

(13) Deisenhofer, J.; Epp, O.; Miki, K.; Huber, R.; Michel, H. *Nature* **1985**, *318* (6047), 618.

(14) Umena, Y.; Kawakami, K.; Shen, J.-R.; Kamiya, N. *Nature* **2011**, *473* (7345), 55.

(15) Kimble, L. K.; Mandelco, L.; Woese, C. R.; Madigan, M. T. *Arch. Microbiol.* **1995**, *163* (4), 259.

(16) Kleinherenbrink, F. A. M.; Ikegami, I.; Hiraishi, A.; Otte, S. C. M.; Amesz, J. *Biochim. Biophys. Acta - Bioenerg.* **1993**, *1142* (1–2), 69.

(17) Olson, J. M.; Blankenship, R. E. In *Photosynthesis Research*; 2004; Vol. 80, pp 373–386.

(18) Khadka, B.; Adeolu, M.; Blankenship, R. E.; Gupta, R. S. *Photosynth. Res.* **2017**, *131* (2), 159.

(19) Lomize, A. L.; Pogozheva, I. D.; Lomize, M. a.; Mosberg, H. I. *Protein Sci.* **2006**, *15* (6), 1318.

(20) Kobayashi, M.; Watanabe, T.; Ikegami, I.; van de Meent, E. J.; Amesz, J. *FEBS Lett.* **1991**, *284* (1), 129.

(21) Chauvet, A.; Sarrou, J.; Lin, S.; Romberger, S. P.; Golbeck, J. H.; Savikhin, S.; Redding, K. E. *Photosynth. Res.* **2013**, *116* (1), 1.

(22) Sarrou, I.; Khan, Z.; Cowgill, J.; Lin, S.; Brune, D.; Romberger, S.; Golbeck, J. H.; Redding, K. E. *Photosynth. Res.* **2012**, *111* (3), 291.

(23) van de Meent, E. J.; Kobayashi, M.; Erkelens, C.; van Veelen, P. A.; Amesz, J.; Watanabe, T. *BBA - Bioenerg.* **1991**, *1058* (3), 356.

(24) Müller, M. G.; Niklas, J.; Lubitz, W.; Holzwarth, A. R. *Biophys. J.* **2003**, *85* (6), 3899.

(25) Müller, M. G.; Slavov, C.; Luthra, R.; Redding, K. E.; Holzwarth, A. R. *Proc. Natl. Acad. Sci. U. S. A.* **2010**, *107* (9), 4123.

(26) Prokhorenko, V. I.; Holzwarth, A. R. *J. Phys. Chem. B* **2000**, *104* (48), 11563.

(27) van Brederode, M. E.; Jones, M. R.; van Mourik, F.; van Stokkum, I. H. M.; van Grondelle, R. *Biochemistry* **1997**, *36* (23), 6855.

(28) van Brederode, M. E.; van Grondelle, R. *FEBS Letters*. 1999, pp 1–7.

(29) Brederode, M. E. van; Mourik, F. van; Stokkum, I. H. M. van; Jones, M. R.; van Brederode, M. E.; van Mourik, F.; van Stokkum, I. H.; Jones, M. R.; van Grondelle, R.; Universiteit, V.; Bank, W.; Kingdom, U. *Proc. Natl. Acad. Sci. U. S. A.* **1999**, *96* (5), 2054.

(30) Ferlez, B.; Cowgill, J.; Dong, W.; Gisriel, C.; Lin, S. S.; Flores, M.; Walters, K.; Cetnar, D.; Redding, K. E. K. E.; Golbeck, J. H. J. H. *Biochemistry* **2016**, *55* (16), 2358.

(31) Nuijs, A. M.; van Dorssen, R. J.; Duysens, L. N. M.; Amesz, J. *Proc. Natl. Acad. Sci. U. S. A.* **1985**, *82* (20), 6865.

(32) Lin, S.; Chiou, H. C.; Kleinherenbrink, F. A.; Blankenship, R. E. *Biophys. J.* **1994**, *66* (2), 437.

(33) Mix, L. J.; Haig, D.; Cavanaugh, C. M. *J. Mol. Evol.* **2005**, *60* (2), 153.

(34) Noguchi, T. *Photosynthesis Research*. 2010, pp 321–331.

(35) Nelson, N.; Yocum, C. F. *Annu. Rev. Plant Biol.* **2006**, *57* (1), 521.

(36) Fromme, R.; Fromme, P.; Grotjohann, I. In *Structural Biology of Membrane Proteins*; Buchanan, S., Grisshammer, R., Eds.; Royal Society of Chemistry, 2006; pp 320–348.

(37) Li, Y.; van der Est, A.; Lucas, M. G.; Ramesh, V. M.; Gu, F.; Petrenko, A.; Lin, S.; Webber, A. N.; Rappaport, F.; Redding, K. *Proc. Natl. Acad. Sci. U. S. A.* **2006**, *103* (7), 2144.

(38) Heinnickel, M.; Agalarov, R.; Svensen, N.; Krebs, C.; Golbeck, J. H.; Park, U. V; Pennsyl, V. *Biochemistry* **2006**, *45* (21), 6756.

(39) Hans, M.; Buckel, W.; Bill, E. *Eur. J. Biochem.* **2000**, *267* (24), 7082.

(40) Romberger, S. P.; Castro, C.; Sun, Y.; Golbeck, J. H. *Photosynth. Res.* **2010**, *104* (2), 293.

(41) Romberger, S. P.; Golbeck, J. H. *Photosynth. Res.* **2012**, *111* (3), 285.

(42) Albert, I.; Rutherford, A. W.; Grav, H.; Kellermann, J.; Michel, H. *Biochemistry* **1998**, *37* (25), 9001.

(43) Kashey, T. S.; Cowgill, J. B.; McConnell, M. D.; Flores, M.; Redding, K. E. *Photosynth. Res.* **2014**, *120* (3), 291.

(44) Sommer, F.; Drepper, F.; Hippler, M. *J. Biol. Chem.* **2002**, *277* (8), 6573.

(45) Sommer, F.; Drepper, F.; Haehnel, W.; Hippler, M. *J. Biol. Chem.* **2004**, *279* (19), 20009.

(46) Gupta, R. S. *Photosynthesis Research*. 2003, pp 173–183.

(47) Sattley, W. M.; Madigan, M. T.; Swingley, W. D.; Cheung, P. C.; Clocksin, K. M.; Conrad, A. L.; Dejesa, L. C.; Honchak, B. M.; Jung, D. O.; Karbach, L. E.; Kurdoglu, A.; Lahiri, S.; Mastrian, S. D.; Page, L. E.; Taylor, H. L.; Wang, Z. T.; Raymond, J.; Chen, M.; Blankenship, R. E.; Touchman, J. W. *J. Bacteriol.* **2008**, *190* (13), 4687.

(48) Gunner, M. R.; Nicholls, A.; Honig, B. *J. Phys. Chem.* **1996**, *100* (10), 4277.

CHAPTER 4

(1) Hohmann-Marriott, M. F.; Blankenship, R. E. *Annu. Rev. Plant Biol.* **2011**, *62*, 515.

(2) Fischer, W. W.; Hemp, J.; Johnson, J. E. *Annu. Rev. Earth Planet. Sci.* **2016**, *44* (1), 647.

(3) Nelson, N.; Yocum, C. F. *Annu. Rev. Plant Biol.* **2006**, *57* (1), 521.

(4) Blankenship, R. E. *Molecular Mechanisms of Photosynthesis*, 2nd ed.; John Wiley & Sons, Ltd.: Southern Gate, Chicester, West Sussex, UK, 2008.

(5) Vassiliev, I. R.; Antonkine, M. L.; Golbeck, J. H. *Biochim. Biophys. Acta - Bioenerg.* **2001**, *1507* (1–3), 139.

(6) Diner, B. A.; Petrouleas, V.; Wendoloski, J. J. *Physiol. Plant.* **1991**, *81* (3), 423.

(7) Cardona, T.; Sedoud, A.; Cox, N.; Rutherford, A. W. *Biochim. Biophys. Acta - Bioenerg.* **2012**, *1817* (1), 26.

(8) Cardona, T. *Photosynthesis Research*. 2015, pp 111–134.

(9) Keeling, P. J. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* **2010**, *365* (1541), 729.

(10) Sadekar, S.; Raymond, J.; Blankenship, R. E. *Mol. Biol. Evol.* **2006**, *23* (11), 2001.

(11) Zeng, Y.; Feng, F.; Medová, H.; Dean, J.; Koblížek, M. *Proc. Natl. Acad. Sci. U. S. A.* **2014**, *111* (21), 7795.

(12) Bryant, D. A.; Frigaard, N.-U. U. *Prokaryotic photosynthesis and phototrophy illuminated*; 2006; Vol. 14, pp 488–496.

(13) Barber, J. *Cold Spring Harb. Symp. Quant. Biol.* **2012**, *77*, 295.

(14) Schubert, W.-D.; Klukas, O.; Saenger, W.; Witt, H. T.; Fromme, P.; Krauß, N. *J. Mol. Biol.* **1998**, *280* (2), 297.

(15) Jordan, P.; Fromme, P.; Witt, H. T.; Klukas, O.; Saenger, W.; Krauss, N. *Nature* **2001**, *411* (6840), 909.

(16) Rémigy, H.-W.; Stahlberg, H.; Fotiadis, D.; Müller, S. A.; Wolpensinger, B.; Engel, A.; Hauska, G.; Tsiotis, G. *J. Mol. Biol.* **1999**, *290* (4), 851.

(17) Heinnickel, M.; Golbeck, J. H. *Photosynth. Res.* **2007**, *92* (1), 35.

(18) Garcia Costas, A. M.; Liu, Z.; Tomsho, L. P.; Schuster, S. C.; Ward, D. M.; Bryant, D. A. *Environ. Microbiol.* **2012**, *14* (1), 177.

(19) Mix, L. J.; Haig, D.; Cavanaugh, C. M. *J. Mol. Evol.* **2005**, *60* (2), 153.

(20) Cardona, T. *Front. Plant Sci.* **2016**, *7* (March), 1.

(21) Rost, B. *Protein Eng. Des. Sel.* **1999**, *12* (2), 85.

(22) Doolittle, R. F. *Of urfs and orfs: a primer on how to analyze derived amino acid sequences*; University Science Books: Mill Valley, CA, 1986.

(23) Battistuzzi, F. U.; Hedges, S. B. *Mol. Biol. Evol.* **2009**, *26* (2), 335.

(24) Khadka, B.; Adeolu, M.; Blankenship, R. E.; Gupta, R. S. *Photosynth. Res.* **2017**, *131* (2), 159.

(25) Gisriel, C.; Sarrou, I.; Ferlez, B. H.; Golbeck, J. H.; Redding, K. E.; Fromme, R. *Science* **2017**, *357* (6355), 1021.

(26) Blankenship, R. E. *Photosynth. Res.* **1992**, *33* (October), 91.

(27) Nitschke, W.; Rutherford, A. W. *Trends Biochem. Sci.* **1991**, *16* (7), 241.

(28) Vermaas, W. F. J. *Photosynth. Res.* **1994**, *41* (1), 285.

(29) Miller, K. R.; Jacob, J. S.; Smith, U.; Kolaczkowski, S.; Bowman, M. K. *Arch. Microbiol.* **1986**, *146* (2), 111.

(30) Pribil, M.; Labs, M.; Leister, D. *Journal of Experimental Botany*. 2014, pp 1955–1972.

(31) Prince, R. C.; Gest, H.; Blankenship, R. E. *BBA - Bioenerg.* **1985**, *810* (3), 377.

(32) Albert, I.; Rutherford, A. W.; Grav, H.; Kellermann, J.; Michel, H. *Biochemistry* **1998**, *37* (25), 9001.

(33) Guergova-Kuras, M.; Boudreaux, B.; Joliot, A.; Joliot, P.; Redding, K. E. *Proc. Natl. Acad. Sci. U. S. A.* **2001**, *98* (8), 4437.

(34) Jagannathan, B.; Golbeck, J. H. *Cellular and Molecular Life Sciences*. 2009, pp 1257–1270.

(35) Büttner, M.; Xie, D. L.; Nelson, H.; Pinther, W.; Hauska, G.; Nelson, N. *Biochim. Biophys. Acta* **1992**, *1101* (2), 154.

(36) Liebl, U.; Mockensturm-Wilson, M.; Trost, J. T.; Brune, D. C.; Blankenship, R. E.; Vermaas, W. F. J. *Proc. Natl. Acad. Sci. U. S. A.* **1993**, *90* (15), 7124.

(37) Bryant, D. A.; Garcia Costas, A. M.; Maresca, J. A.; Gomez Maqueo Chew, A.; Klatt, C. G.; Bateson, M. M.; Tallon, L. J.; Hostetler, J.; Nelson, W. C.; Heidelberg, J. F.; Ward, D. M. *Science.* **2007**, *317* (5837), 523.

(38) Chitnis, P. R.; Xu, Q.; Chitnis, V. P.; Nechushtai, R. *Photosynthesis Research*.

1995, pp 23–40.

(39) Vos, M. H.; Klaassen, H. E.; van Gorkom, H. J. *BBA - Bioenerg.* **1989**, *973* (2), 163.

(40) Nitschke, W.; Liebl, U.; Matsuura, K.; Kramer, D. M. *Biochemistry* **1995**, *34* (37), 11831.

(41) Hervás, M.; Navarro, J. a; Díaz, a; Bottin, H.; De la Rosa, M. a. *Biochemistry* **1995**, *34* (36), 11321.

(42) Hippler, M.; Drepper, F.; Farah, J.; Rochaix, J. D. *Biochemistry* **1997**, *36* (21), 6343.

(43) Xiong, J.; Inoue, K.; Bauer, C. E. *Proc. Natl. Acad. Sci. U. S. A.* **1998**, *95* (25), 14851.

(44) Bryantseva, I. A.; Gorlenko, V. M.; Kompantseva, E. I.; Achenbach, L. A.; Madigan, M. T. *Arch. Microbiol.* **1999**, *172* (3), 167.

(45) Fromme, P.; Schubert, W. D.; Krauß, N. *BBA - Bioenergetics*. 1994, pp 99–105.

(46) Heinnickel, M.; Shen, G.; Agalarov, R.; Golbeck, J. H. *Biochemistry* **2005**, *44* (29), 9950.

(47) Heinnickel, M.; Shen, G.; Golbeck, J. H. *Biochemistry* **2007**, *46* (9), 2530.

(48) Sarrou, I.; Khan, Z.; Cowgill, J.; Lin, S.; Brune, D. C.; Romberger, S.; Golbeck, J. H.; Redding, K. E. *Photosynth. Res.* **2012**, *111* (3), 291.

(49) Ferlez, B.; Cowgill, J.; Dong, W.; Gisriel, C.; Lin, S. S.; Flores, M.; Walters, K.; Cetnar, D.; Redding, K. E. K. E.; Golbeck, J. H. J. H. *Biochemistry* **2016**, *55* (16), 2358.

(50) Jagannathan, B.; Golbeck, J. H. *Biochemistry* **2009**, *48* (23), 5405.

(51) Jagannathan, B.; Dekat, S.; Golbeck, J. H.; Lakshmi, K. V. *Biochemistry* **2010**, *49* (11), 2398.

(52) Fromme, P.; Jordan, P.; Krauß, N. *Biochimica et Biophysica Acta - Bioenergetics*. 2001, pp 5–31.

(53) Okamura, M. Y.; Feher, G. *Annu. Rev. Biochem.* **1992**, *61*, 861.

(54) Okamura, M. Y.; Paddock, M. L.; Graige, M. S.; Feher, G. *Biochim. Biophys. Acta* **2000**, *1458* (1), 148.

(55) Graige, M. S.; Paddock, M. L.; Bruce, J. M.; Feher, G.; Okamura, M. Y. *J. Am. Chem. Soc.* **1996**, *118* (38), 9005.

(56) Brok, M.; Vasmel, H.; Horikx, J. T. G.; Hoff, A. J. *FEBS Lett.* **1986**, *194* (2), 322.

(57) Brettel, K.; Leibl, W.; Liebl, U. *Biochimica et Biophysica Acta - Bioenergetics*. 1998, pp 175–181.

(58) Lin, S.; Chiou, H. C.; Blankenship, R. E. *Biochemistry* **1995**, *34* (39), 12761.

(59) van Der Est, A.; Hager-Braun, C.; Leibl, W.; Hauska, G.; Stehlik, D. *Biochim. Biophys. Acta - Bioenerg.* **1998**, *1409* (2), 87.

(60) Trost, J. T.; Brune, D. C.; Blankenship, R. E. *Photosynth. Res.* **1992**, *32* (1), 11.

(61) Miyamoto, R.; Mino, H.; Kondo, T.; Itoh, S.; Oh-oka, H. *Biochemistry* **2008**, *47* (15), 4386.

(62) Kondo, T.; Itoh, S.; Matsuoka, M.; Azai, C.; Oh-Oka, H. *J. Phys. Chem. B* **2015**, *119* (27), 8480.

(63) Kleinherenbrink, F. A. M.; Ikegami, I.; Hiraishi, A.; Otte, S. C. M.; Amesz, J. *Biochim. Biophys. Acta - Bioenerg.* **1993**, *1142* (1–2), 69.

(64) Trost, J. T.; Blankenship, R. E. *Biochemistry* **1989**, *28* (26), 9898.

(65) Kashey, T. Recombinant electron donors and acceptors to and from reaction center particles, and light dependent menaquinone reduction in isolated membranes of *Heliobacterium modesticaldum*, Arizona State University, 2015.

(66) Nitschke, W.; Mattioli, T.; Rutherford, A. W. In *Origin and Evolution of Biological Energy Conservation*; Baltscheffsky, H., Ed.; VCH: New York, 1996; pp 177–203.

(67) Allen, J. F. *FEBS Lett.* **2005**, *579* (5), 963.

(68) McConnell, M. D.; Cowgill, J. B.; Baker, P. L.; Rappaport, F.; Redding, K. E. *Biochemistry* **2011**, *50* (51), 11034.

(69) Hemm, M. R.; Paul, B. J.; Schneider, T. D.; Storz, G.; Rudd, K. E. *Mol. Microbiol.* **2008**, *70* (6), 1487.

(70) Hansson, A.; Amann, K.; Zygadlo, A.; Meurer, J.; Scheller, H. V; Jensen, P. E. *FEBS J.* **2007**, *274* (7), 1734.

(71) Shi, T.; Bibby, T. S.; Jiang, L.; Irwin, A. J.; Falkowski, P. G. *Mol. Biol. Evol.* **2005**, *22* (11), 2179.

(72) Heathcote, P.; Fyfe, P. K.; Jones, M. R. *Trends in Biochemical Sciences*. 2002, pp 79–87.

(73) Olson, J. M. *Biosystems* **1981**, *14* (1), 89.

(74) Nitschke, W.; Rutherford, A. W. *Trends Biochem. Sci.* **1991**, *16* (7), 241.

(75) Rutherford, A. W.; Mattiolo, T. A.; Nitschke, W. In *Origin and evolution of biological energy conversion*; Baltscheffsky, H., Ed.; New York, 1996; pp 177–204.

(76) Olson, J. M.; Blankenship, R. E. In *Photosynthesis Research*; 2004; Vol. 80, pp 373–386.

(77) Neerken, S.; Amesz, J. *Biochim. Biophys. Acta - Bioenerg.* **2001**, *1507* (1–3), 278.

(78) Zouni, A.; Witt, H.-T.; Kern, J.; Fromme, P.; Krauß, N.; Saenger, W.; Orth, P. *Nature* **2001**, *409* (6821), 739.

(79) Deisenhofer, J.; Epp, O.; Miki, K.; Huber, R.; Michel, H. *Nature* **1985**, *318* (6047), 618.

(80) Lince, M. T.; Vermaas, W. *Eur. J. Biochem.* **1998**, *256* (3), 595.

(81) Schelvis, J. P. M.; van Noort, P. I.; Aartsma, T. J.; van Gorkom, H. J. *Biochim. Biophys. Acta - Bioenerg.* **1994**, *1184* (2–3), 242.

(82) Schweitzer, R. H.; Brudvig, G. W. *Biochemistry* **1997**, *36* (38), 11351.

(83) Schweitzer, R. H.; Melkozernov, A. N.; Blankenship, R. E.; Brudvig, G. W. *J. Phys. Chem. B* **1998**, *102* (42), 8320.

(84) Akiyama, M.; Miyashita, H.; Kise, H.; Watanabe, T.; Mimuro, M.; Miyachi, S.; Kobayashi, M. *Photosynth. Res.* **2002**, *74* (2), 97.

(85) Ohashi, S.; Iemura, T.; Okada, N.; Itoh, S.; Furukawa, H.; Okuda, M.; Ohnishi-Kameyama, M.; Ogawa, T.; Miyashita, H.; Watanabe, T.; Itoh, S.; Oh-oka, H.; Inoue, K.; Kobayashi, M. *Photosynth. Res.* **2010**, *104* (2), 305.

(86) Prince, R. C.; Gest, H.; Blankenship, R. E. *BBA - Bioenerg.* **1985**, *810* (3), 377.

(87) Kobayashi, M.; van de Meent, E. J.; Oh-oka, H.; Inoue, K.; Itoh, S.; Amesz, J. In *Research in Photosynthesis*; Murata, N., Ed.; Kluwer Academic Publishers: Dordrecht, The Netherlands, 1992; pp 393–396.

(88) Poluektov, O. G.; Paschenko, S. V; Utschig, L. M.; Lakshmi, K. V; Thurnauer, M. C. *J. Am. Chem. Soc.* **2005**, *127* (34), 11910.

(89) Santabarbara, S.; Heathcote, P.; Evans, M. C. W. *Biochimica et Biophysica Acta - Bioenergetics*. 2005, pp 283–310.

(90) Muhiuddin, I. P.; Heathcote, P.; Carter, S.; Purton, S.; Rigby, S. E. J.; Evans, M. C. W. *FEBS Lett.* **2001**, *503* (1), 56.

(91) Webber, A. N.; Lubitz, W. *Biochimica et Biophysica Acta - Bioenergetics*. 2001, pp 61–79.

(92) Li, Y.; Lucas, M. G.; Konovalova, T.; Abbott, B.; MacMillan, F.; Petrenko, A.; Sivakumar, V.; Wang, R.; Hastings, G.; Gu, F.; Van Tol, J.; Brunel, L. C.; Timkovich, R.; Rappaport, F.; Redding, K. *Biochemistry* **2004**, *43* (39), 12634.

(93) Chen, M.; Telfer, A.; Lin, S.; Pascal, A.; Larkum, A. W. D.; Barber, J.; Blankenship, R. E. *Photochem. Photobiol. Sci.* **2005**, *4* (12), 1060.

(94) Tomo, T.; Okubo, T.; Akimoto, S.; Yokono, M.; Miyashita, H.; Tsuchiya, T.; Noguchi, T.; Mimuro, M. *Proc. Natl. Acad. Sci. U. S. A.* **2007**, *104* (17), 7283.

(95) Pan, J.; Saer, R. G.; Lin, S.; Beatty, J. T.; Woodbury, N. W. *Biochemistry* **2016**, acs. biochem.6b00317.

(96) Kirmaier, C.; Laporte, L.; Schenck, C. C.; Holten, D. *J. Phys. Chem.* **1995**, *99* (21), 8910.

(97) Dibrova, D. V; Cherepanov, D. A.; Galperin, M. Y.; Skulachev, V. P.; Mulkidjanian, A. Y. *Biochim. Biophys. Acta - Bioenerg.* **2013**, *1827* (11–12), 1407.

(98) Furbacher, P. N.; Tae, G. S.; Cramer, W. A. In *Origin and Evolution of Biological Energy Conservation*; Baltscheffsky, H., Ed.; Wiley-VCH: New York, 1996; pp 221–253.

(99) Bak, D. W.; Elliott, S. J. *Curr. Opin. Chem. Biol.* **2014**, *19* (Supplement C), 50.

(100) Okamura, M. . Y.; Paddock, M. . L.; Graige, M. . S.; Feher, G. *Biochim. Biophys. Acta - Bioenerg.* **2000**, *1458* (1), 148.

(101) Romberger, S. P.; Golbeck, J. H. *Photosynth. Res.* **2012**, *111* (3), 285.

(102) Cardona, T.; Sanchez-Baracaldo, P.; Rutherford, A. W.; Larkum, A. **2017**.

(103) Chauvet, A.; Sarrou, J.; Lin, S.; Romberger, S. P.; Golbeck, J. H.; Savikhin, S.; Redding, K. E. *Photosynth. Res.* **2013**, *116* (1), 1.

(104) Warren, P. V.; Golbeck, J. H.; Warden, J. T. *Biochemistry* **1993**, *32* (3), 849.

(105) Rutherford, A. W.; Osyczka, A.; Rappaport, F. *FEBS Letters*. 2012, pp 603–616.

(106) Berman, H. M.; Westbrook, J.; Zukang, F.; Gilliland, G.; Bhat, T. N.; Weissig, H.; Shindyalov, I. N.; Bourne, P. E. *Nucleic Acids Res.* **2000**, *28* (1), 235.

(107) DeLano, W. L. *Schrödinger LLC* **2014**, http://www.pymol.org.

(108) The Uniprot Consortium. *Nucleic Acids Res.* **2017**, *45* (D1), D158.

(109) Pei, J.; Kim, B.-H. H.; Grishin, N. V. *Nucleic Acids Res.* **2008**, *36* (7), 2295.

(110) Krissinel, E.; Henrick, K. In *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*; 2005; Vol. 3695 LNBI, pp 67–78.

(111) Sievers, F.; Wilm, A.; Dineen, D.; Gibson, T. J.; Karplus, K.; Li, W.; Lopez, R.; McWilliam, H.; Remmert, M.; Söding, J.; Thompson, J. D.; Higgins, D. G. *Mol. Syst. Biol.* **2011**, *7* (1).

(112) Goujon, M.; McWilliam, H.; Li, W.; Valentin, F.; Squizzato, S.; Paern, J.; Lopez, R. *Nucleic Acids Res.* **2010**, *38* (SUPPL. 2).

(113) Kumar, S.; Stecher, G.; Tamura, K. *Mol. Biol. Evol.* **2016**, msw054.

(114) Adachi, J.; Waddell, P. J.; Martin, W.; Hasegawa, M. *J. Mol. Evol.* **2000**, *50* (4), 348.

(115) Whelan, S.; Goldman, N. *Mol. Biol. Evol.* **2001**, *18* (5), 691.

(116) Le, S. Q.; Gascuel, O. *Mol. Biol. Evol.* **2008**, *25* (7), 1307.

(117) Dimmic, M. W.; Rest, J. S.; Mindell, D. P.; Goldstein, R. A. *J. Mol. Evol.* **2002**, *55* (1), 65.

CHAPTER 5

(1) Das, D.; Veziroglu, T. N. *Int. J. Hydrogen Energy* **2001**, *26* (1), 13.

(2) Levin, D. B.; Pitt, L.; Love, M. *Int. J. Hydrogen Energy* **2004**, *29* (2), 173.

(3) Benemann, J. R. *Int. J. Hydrogen Energy* **1997**, *22* (10–11), 979.

(4) Sarkar, S.; Pandey, K. D.; Kashyap, A. K. *Int. J. Hydrogen Energy* **1992**, *17* (9), 689.

(5) Lambert, G. R.; Smith, G. D. *Arch. Biochem. Biophys.* **1980**, *205* (1), 36.

(6) Tel-Or, E.; Luijk, L. W.; Packer, L. *FEBS Lett.* **1977**, *78* (1), 49.

(7) Bothe, H.; Distler, E.; Eisbrenner, G. *Biochimie* **1978**, *60* (3), 277.

(8) Kumar, N.; Das, D. *Process Biochem.* **2000**, *35* (6), 589.

(9) Tanisho, S.; Wakao, N.; Kosako, Y. *J. Chem. Eng. Jpn.* **1983**, *16* (6), 529.

(10) Tanisho, S.; Suzuki, Y.; Wakao, N. *Int. J. Hydrogen Energy* **1987**, *12* (9), 623.

(11) Brosseau, J. D.; Zajic, J. E. *Int. J. Hydrogen Energy* **1982**, *7* (8), 623.

(12) Luque, R.; Lin, C. S.; Wilson, K.; Clark, J. Elsevier Inc., 2010; pp 313–314.

(13) Vignais, P. M.; Billoud, B.; Meyer, J. *FEMS Microbiology Reviews*. 2001, pp 455–501.

(14) Meyer, J. *Cellular and Molecular Life Sciences*. 2007, pp 1063–1084.

(15) Vignais, P. M.; Billoud, B. *Chem. Rev.* **2007**, *107* (10), 4206.

(16) Posewitz, M. C.; Mulder, D. W.; Peters, J. W. *Curr. Chem. Biol.* **2008**, *2* (2), 178.

(17) Mulder, D. W.; Shepard, E. M.; Meuser, J. E.; Joshi, N.; King, P. W.; Posewitz, M. C.; Broderick, J. B.; Peters, J. W. *Structure*. 2011, pp 1038–1052.

(18) Canfield, D. E.; Glazer, A. N.; Falkowski, P. G. *Science.* **2010**, *330* (6001), 192.

(19) Thamdrup, B. *Annu. Rev. Ecol. Evol. Syst.* **2012**, *43* (1), 407.

(20) Gruber, N.; Galloway, J. N. *Nature* **2008**, *451* (7176), 293.

(21) Peters, J. W.; Szilagyi, R. K. *Current Opinion in Chemical Biology*. 2006, pp 101–108.

(22) Hoffman, B. M.; Lukoyanov, D.; Yang, Z.-Y.; Dean, D. R.; Seefeldt, L. C. *Chem. Rev.* **2014**, *114* (8), 4041.

(23) Burgess, B. K.; Lowe, D. J. *Chem. Rev.* **1996**, *96* (7), 2983.

(24) Bryantseva, I. A.; Gorlenko, V. M.; Kompantseva, E. I.; Achenbach, L. A.; Madigan, M. T. *Arch. Microbiol.* **1999**, *172* (3), 167.

(25) Tang, K.-H.; Yue, H.; Blankenship, R. E. *BMC Microbiol.* **2010**, *10*, 150.

(26) Kimble, L. K.; Madigan, M. T. *Arch. Microbiol.* **1992**, *158*, 155.

(27) Ludden, P. W.; Roberts, G. P.; Lowery, R. G.; Fitzmaurice, W. P.; Saari, L. L.; Lehman, L.; Lies, D.; Woehle, D.; Wirt, H.; Murrell, S. A. *Nitrogen Fixat. hundred years after Proc. 7th Int. Congr. N [Triple-bond] Nitrogen Fixat. Koln (Cologne), F.R.G., March 13-20, 1980 / Ed. by H. Bothe, F.J. Bruijn W.E. Newt.* **1988**.

(28) Kimble, L. K.; Mandelco, L.; Woese, C. R.; Madigan, M. T. *Arch. Microbiol.* **1995**, *163* (4), 259.

(29) Xiong, J.; Inoue, K.; Bauer, C. E. *Proc. Natl. Acad. Sci. U. S. A.* **1998**, *95* (25), 14851.

(30) Laureanti, J. A. Bioinspired Electrocatalytic Hydrogen Production: Synthetic and Biological Approaches, Arizona State University, 2017.

(31) Nevin, K. P.; Hensley, S. A.; Franks, A. E.; Summers, Z. M.; Ou, J.; Woodard, T. L.; Snoeyenbos-West, O. L.; Lovley, D. R. *Appl. Environ. Microbiol.* **2011**, *77* (9), 2882.

(32) Fuller, R. C.; Sprague, S. G.; Gest, H.; Blankenship, R. E. *FEBS Lett.* **1985**, *182* (2), 345.

(33) Miller, K. R.; Jacob, J. S.; Smith, U.; Kolaczkowski, S.; Bowman, M. K. *Arch. Microbiol.* **1986**, *146* (2), 111.

(34) Kudryashev, M.; Aktoudianaki, A.; Dedoglou, D.; Stahlberg, H.; Tsiotis, G. *Biochim. Biophys. Acta - Bioenerg.* **2014**, *1837* (10), 1635.

(35) Okubo, Y.; Futamata, H.; Hiraishi, A. *Appl. Environ. Microbiol.* **2006**, *72* (9), 6225.

APPENDIX A

SUPPLEMENTARY INFORMATION FOR CHAPTER 3

Supplementary Information

Crystallization and data collection. Heliobacterial reaction centers (HbRCs) were solubilized from heliobacterial membranes with β-dodecylmaltoside (β-DDM) and purified as described [1]. Crystallization experiments were conducted under anoxic conditions (~97% N_2, 3% H_2) using hanging-drop Linbro plates. The protein was crystallized by evaporative diffusion in the C_2 space group (a = 131 Å, b=90 Å, c=112 Å, α=90° β=108.15° γ=90°) at 20 °C in a solution containing 0.3 mM HbRC, 100 mM Tris-HCl (pH 8.2), 600 mM NH_4Cl, 0.02% β-DDM, 1% *n*-heptyl-β-D-glucopyranoside, and 20% PEG500. Protein crystals were harvested in a solution containing an identical composition to which they were crystallized, except the PEG500 concentration was raised to 35%. Each crystal was flash-cooled in liquid N_2 and stored at 77 K until measurement. Native x-ray diffraction data sets were collected at ALS beamline 8.2.1 of the Advanced Light Source (ALS, Berkeley, CA) and beamline 19-BM of the Advanced Photon Source (APS, Argonne, IL).

Structure determination. Native data sets at 2.2-Å resolution were collected under cryogenic conditions. Information on data collection, evaluation, and refinement is provided in Table S1. The crystals had a typical size of 150 x 200 x 300 μm and belonged to the C_2 space group. Initial phases were obtained by molecular replacement [2] with a PSI homology model. The homology model of PSI used for initial phases was created using a minimized alteration of the known structure of chain A of PSI [Protein Data Bank (PDB) ID: 1JB0]. All side-chains were changed to Ala residues and looping regions were removed, leaving the model with only 11 simple TMH. In addition, only the porphyrin

rings of the chlorophylls expected to be involved in ET and the [4Fe-4S] cluster were included in the model.

The initial phases were improved by using SHELXE [3,4]. Initial electron density maps displayed multiple helical shapes into which the simple helices were adjusted. Solvent flattening sharpened the resultant electron density map. The program Buccaneer [5] was able to assign ~80% of the side chains. During the process of refinement, electron density for other features of the structure was resolved, allowing the sequence of PshA to replace the simple model.

Further refinement with ARP/wARP [6] and phenix.refine [7] enabled the addition of structural details, including cofactors, further improving the model as judged by the decrease in R_{work} and R_{free}. In total, 600 of the 609 amino acid residues of PshA were assigned. The first eight N-terminal residues and the last C-terminal residue were not identifiable in the electron density. In total, 30 (B)Chls were identified per monomer, most with a complete farnesyl tail. A single [4Fe-4S] cluster is jointly coordinated between the two monomers by the two PshA subunits.

During refinement, the data converged to R_{work} and R_{free} values of 15.8% and 19.1%, respectively. Interestingly, only one PshA and one PshX are in the asymmetric unit and the [4Fe-4S] cluster is shared between two adjacent asymmetric units. This is a unique feature among all RC structures and reflects the high symmetry within the HbRC. The biologically relevant assembly, composed of two PshA and two PshX subunits with their associated cofactors, was constructed to visualize the HbRC (Fig. 1). Apparent in the electron density is a pair of dodecyl-maltosides, the major detergent used in

purification and crystallization of the HbRC. They are at the crystal contact between HbRCs, with the maltose hydroxyls and associated waters forming a hydrogen-bond network that links a PshA N-terminal loop (backbone oxygens of residues Pro56, Leu80, Ala82, and the side chain of Asp 79) of one HbRC and to a PshA C-terminal loop (backbone oxygens of Gln566, Gln567, and Val568) of another HbRC.

Identification of PshX. The high resolution of the diffraction data allowed identification of some of the side chains of the two extra TMH found in the electron density map. Based on the presumptive partial sequence, a BLAST search was conducted on the annotated genome of *Heliobacterium modesticaldum* [8]. The top (and only convincing) hit was HM1_0821, a hypothetical open reading frame potentially encoding a 40-residue polypeptide (MWKKGGVRPMEMYSPTFNVAHILAFFFLFLHIPFYFVLKD). A low-MW band migrating just after the front was observed by SDS-PAGE (Fig. S1). Mass spectrometry (performed at UCLA by the group of Julian Whitelegge) of the low-weight polypeptide purified by SDS-PAGE provided an estimate of its mass as 3,827 Da. This is consistent with a 31-residue peptide beginning at the second predicted Met residue, which would be formylated (i.e. formyl-MEMYSPTFNVAHILAFFFLFLHIPFYFVLKD). Thus, this peptide would be the result of translation initiation at the second Met codon.

Quantitation of menaquinone content. The menaquinone (MQ) content of the HbRC preparation used for crystallization was assessed by high-performance liquid chromatography (HPLC), as described previously [9]. Briefly, the HbRC solution was extracted with acetone and pigments were analyzed by reverse-phase HPLC (C18

column, black line in Fig. S4). To assess the sensitivity of the technique, a small amount of MQ-9 (Santa Cruz Biotechnology, Inc.) was added to an extract and analyzed in the same way (red line in Fig. S4).

Estimation of the hydrophobic thickness. The average thickness of the hydrophobic region of the membrane in which an integral membrane protein resides was estimated using the Orientations of Proteins in Membranes database (OMP) and Positioning of Proteins in Membranes server (PPM), which is a molecular dynamics algorithm that optimizes the free energy of protein transfer from water to a lipid bilayer [10]. This tool was utilized to estimate and compare the hydrophobic thickness of the HbRC (PDB ID: 5V8K), PSI (PDB ID: 1JB0), PSII (PDB ID: 3WU2), and the PbRC (PDB ID: 1PRC).

Structural comparison. The antenna domains of type I RCs were defined as the first 6 TMH of the core subunits (PshA, PsaA, and PsaB) and the antenna domain of type II RCs were defined as the 6 TMH from CP43 and CP47 (PsbC and PsbB, respectively). The ET domains of type I RCs were defined as the last 5 TMH and the ET domains of type II RCs were defined as all 5 TMH from the core subunits (PsbA and PsbD for PSII, L and M for the PbRC). The specific residues used for each TMH are listed in Tables S3 and S4. To quantify TMH structural similarity, the TMH of each domain were isolated from the HbRC (PDB ID: 5V8K), PSI (PDB ID: 1JB0), PSII (PDB ID: 3WU2), and the PbRC (PDB ID: 1PRC), and the domains were superimposed onto one-another using the Pymol software [11]. Deviations in the structures were reported as root mean square

deviation (RMSD) values. To quantify the positional similarity of cofactors, all 11 TMH of the core dimer were superimposed onto those of PSI.

Homology modeling and surface electrostatics. A model of PshB1 ferredoxin and cyt c_{553} were created using the Swiss Model software [12]. These were based on sequence identities from published structures (PDB IDs: 1DUR and 1H31, respectively).

Electrostatic surface analysis was performed on the HbRC structure, model of PshB1, and model of cyt c_{553}, by using the APBS server [13] to create surface electrostatics maps of each protein.

Figures

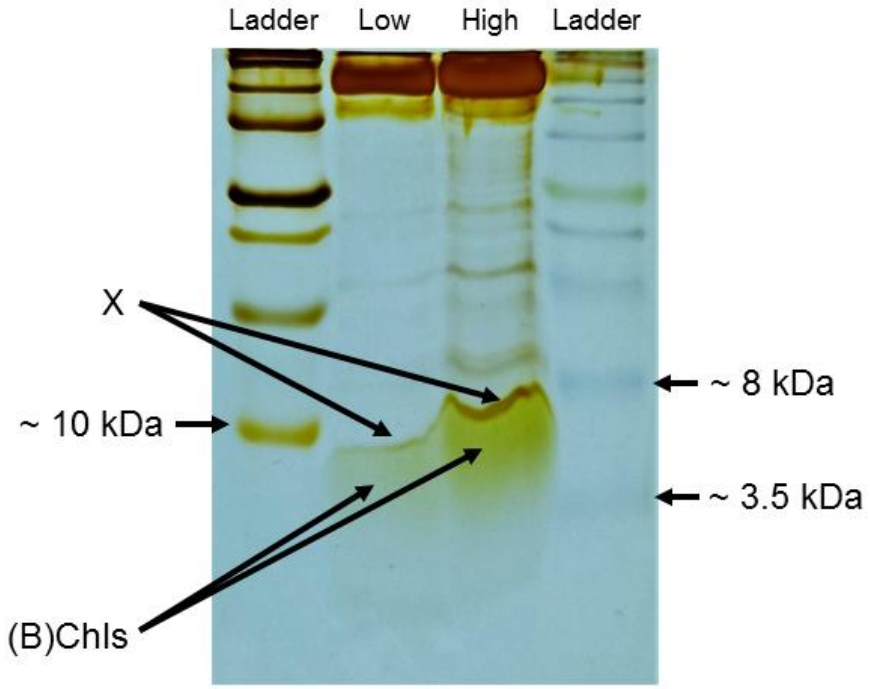
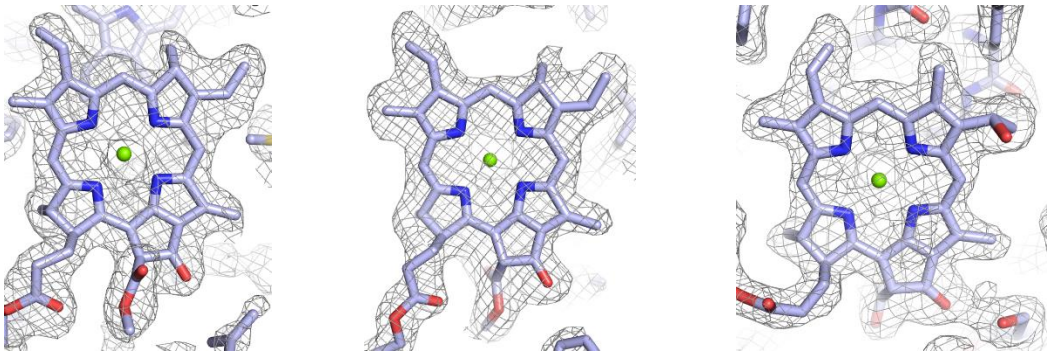


Figure A1 SDS-PAGE of the HbRC preparation. HbRC loaded onto a 20% Tris-Gly gel specialized for resolving low MW polypeptides. Two amounts of HbRC were loaded: Low (2.7 μg protein) and High (9.1 μg protein). Arrows denote the presence of (B)Chls running at the front and a low MW band marked "X". The lowest MW species in the ladders are labeled.

A B C



D

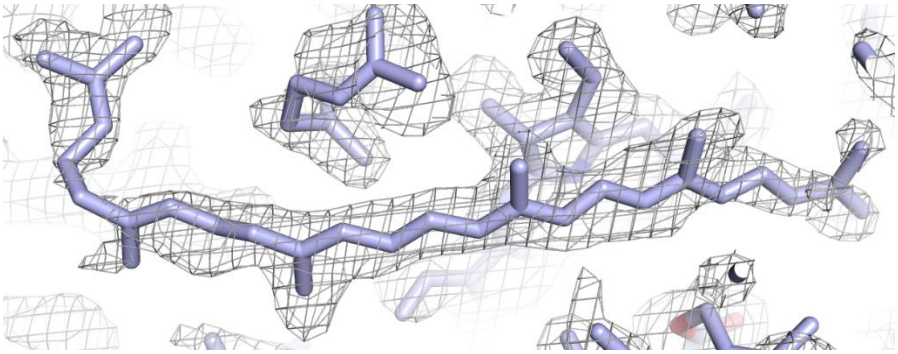


Figure A2 Omit maps of ligands added into the PDB. Simulated-annealing composite OMIT map contoured at 1.0σ containing models of BChl g' (**A**), BChl g (**B**), 8^1-OH-Chl a_F (**C**) and 4,4'-diaponeurosporene (**D**).

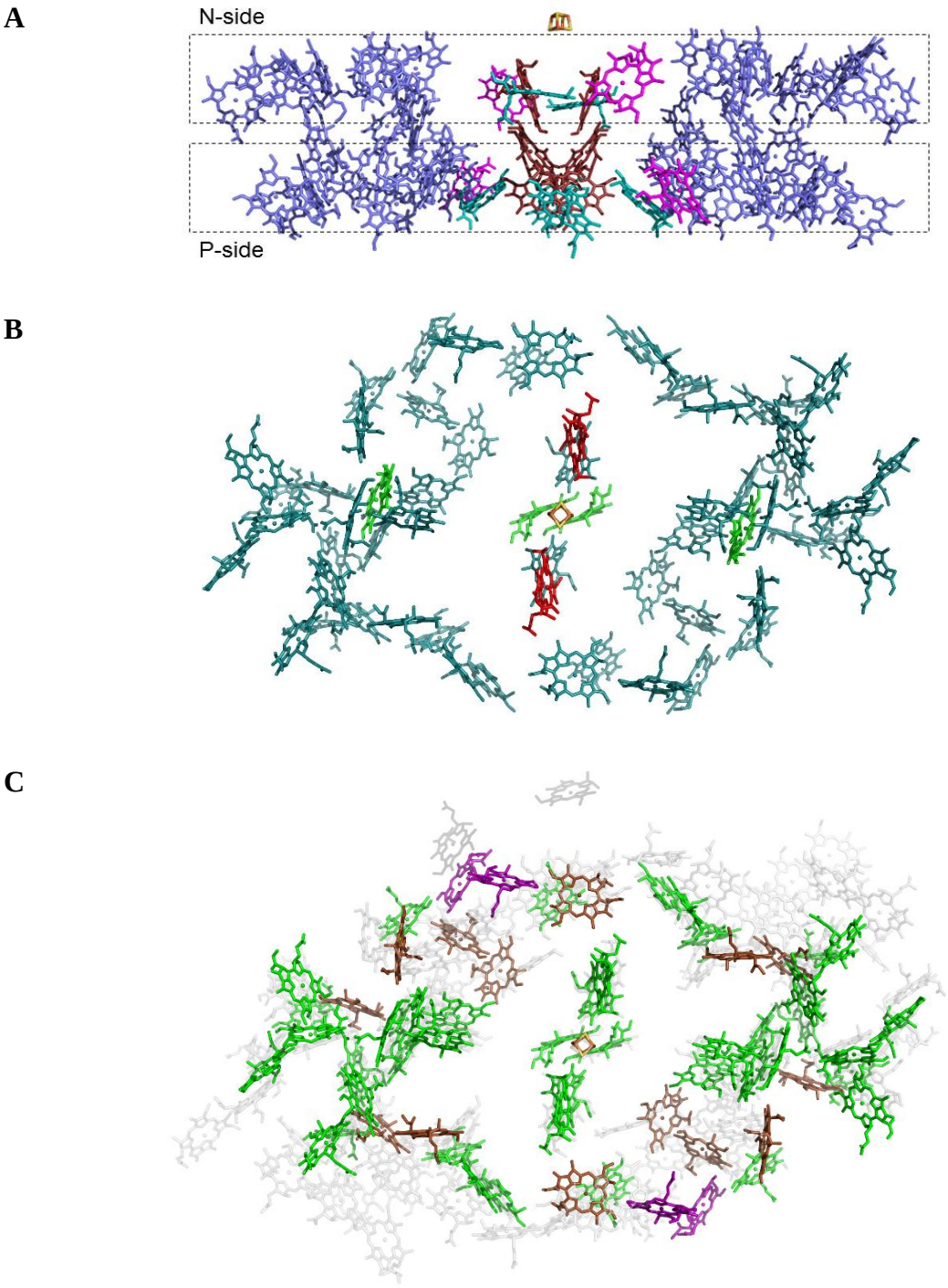


Figure A3 (B)Chl cofactors of the HbRC and comparison with PSI. (**A**) Membrane plane view of the (B)Chls associated with the HbRC. The bulk antenna associated with PshA

are colored in blue, antenna (B)Chls associated with PshX are colored in magenta, BChl g that link the bulk antenna to the ET chain are colored in teal, and the ET chain (B)Chls are colored in red. (**B**) N-side view of the (B)Chls bound to the $PshA_2$ homodimer, colored by chemical identity: BChl g (teal), BChl g' (green), 8^1-OH-Chl a_F (red). (**C**) N-side view of the (B)Chl arrangement of the HbRC and its superposition with the (B)Chls associated with PsaA (light gray, left side), PsaB (light gray, right side), and PsaJ (dark gray, top) in PSI (PDB ID: 1JB0). The (B)Chls of the HbRC that are in the same position as a PSI (B)Chl are colored green, while those that have no analog in PSI are colored brown (PshA) or purple (PshX). All (B)Chl tails have been truncated for clarity.

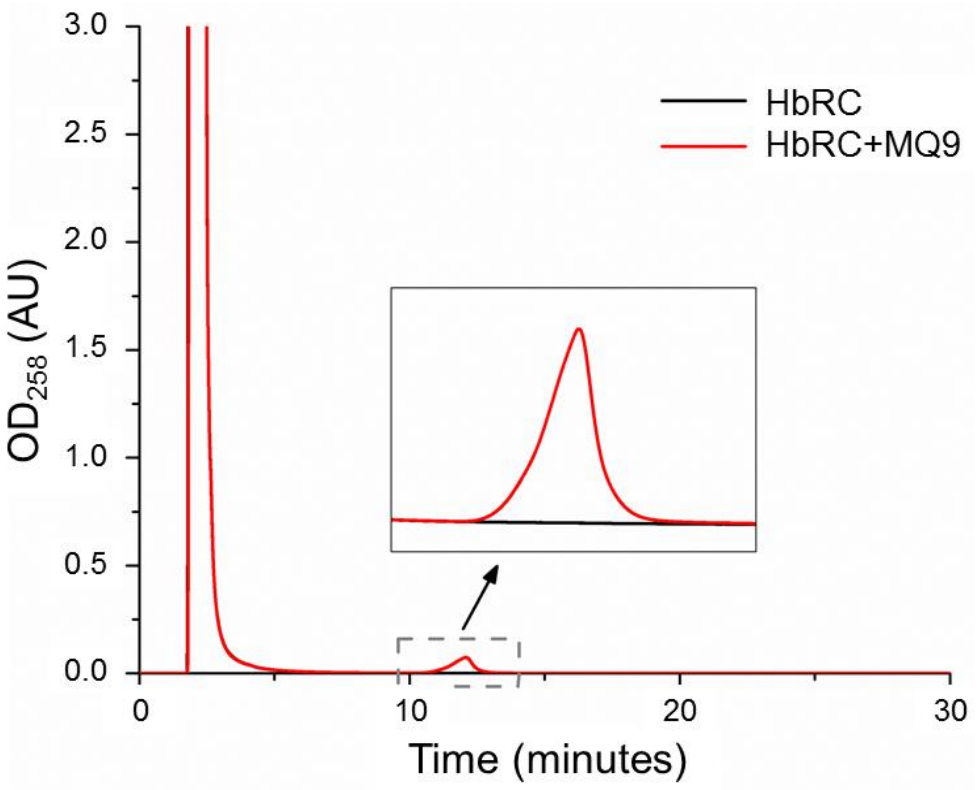


Figure A4 Chromatograms of acetone extracts of HbRC. Acetone extracts of HbRC without (black line) and with (red line) the addition of MQ9 at a molar ratio of 1 MQ9 to 9 HbRC. Acetone extracts were chromatographed on a C18 column as described previously [9]. The inset shows a magnification of the MQ9 peak region in the two chromatograms to demonstrate the sensitivity of the technique. The early peak (~2 minutes) is the injection peak due to acetone and (B)Chl absorption.

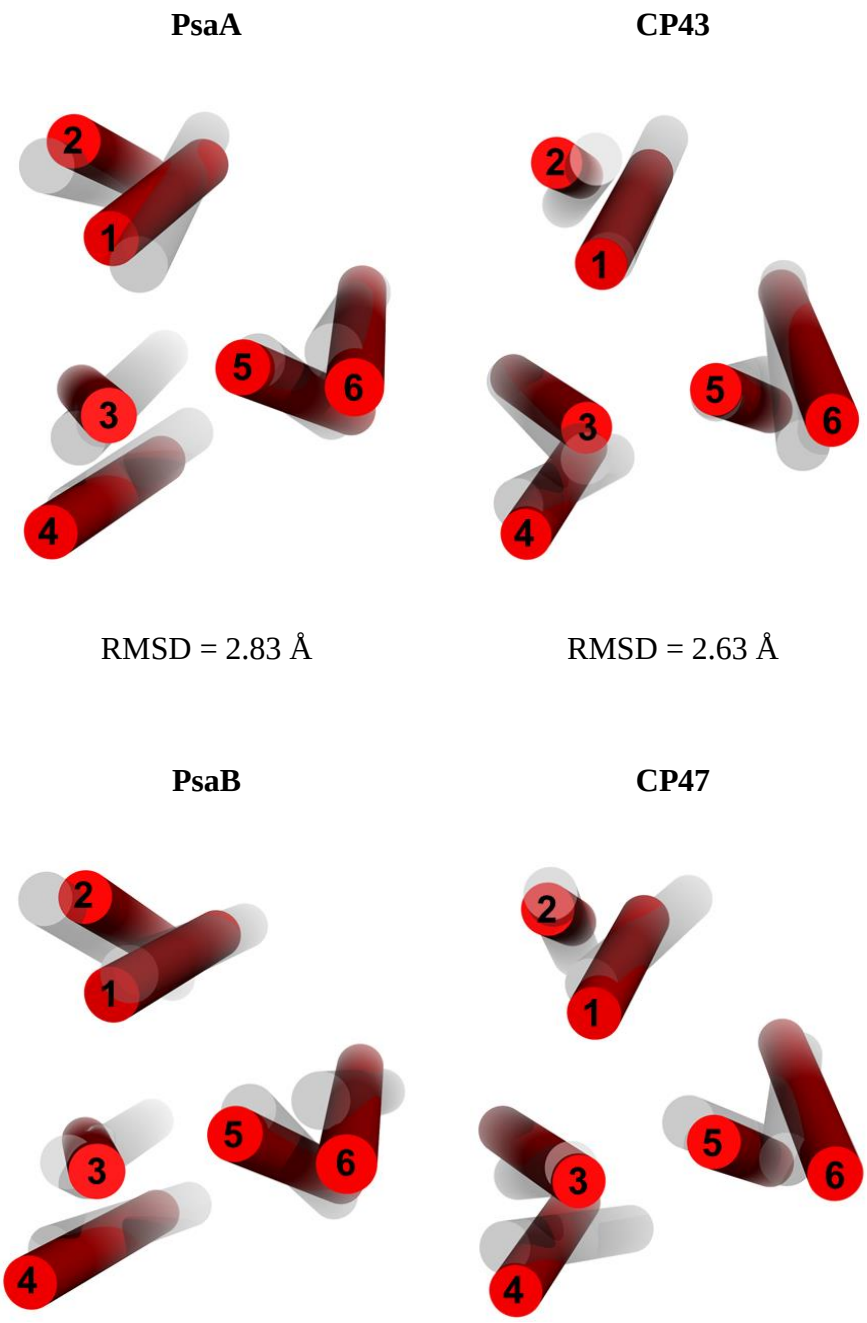


Figure A5 Superposition of antenna domains. Superposition of TMH from the antenna domain of PshA (red) with the analogous TMH from

PsaA, PsaB, CP43, and CP47. Calculated deviations as RMSD values are also shown for each pair.

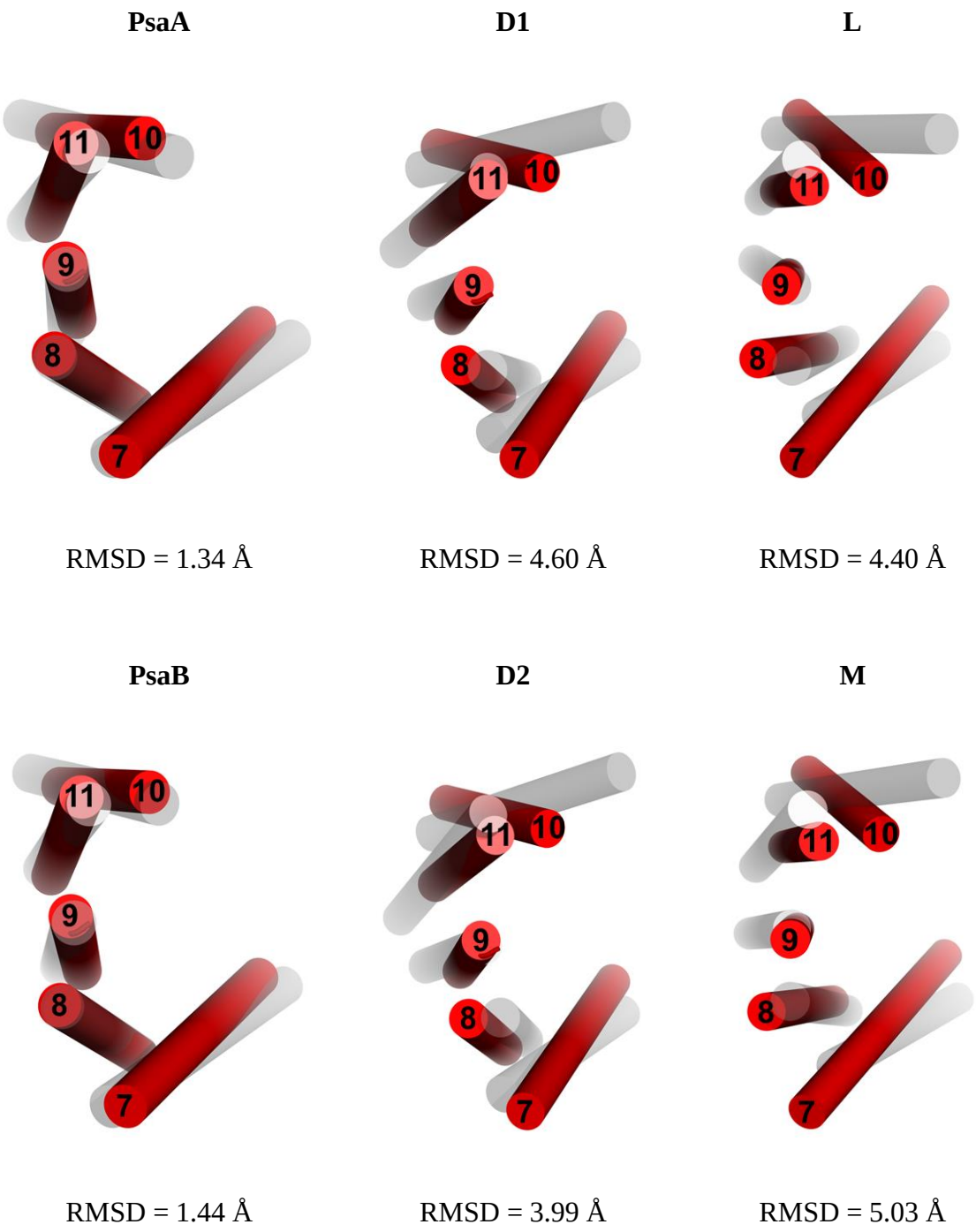


Figure A6 Superposition of ET domains**.** Superposition of TMH from the ET domain of PshA (red) with the analogous TMH from PsaA, PsaB, D1, D2, L, and M, (transparent

grey, PDB IDs: 1JB0, 1PRC, and 3I4D). Calculated deviations as RMSD values are also shown for each pair.

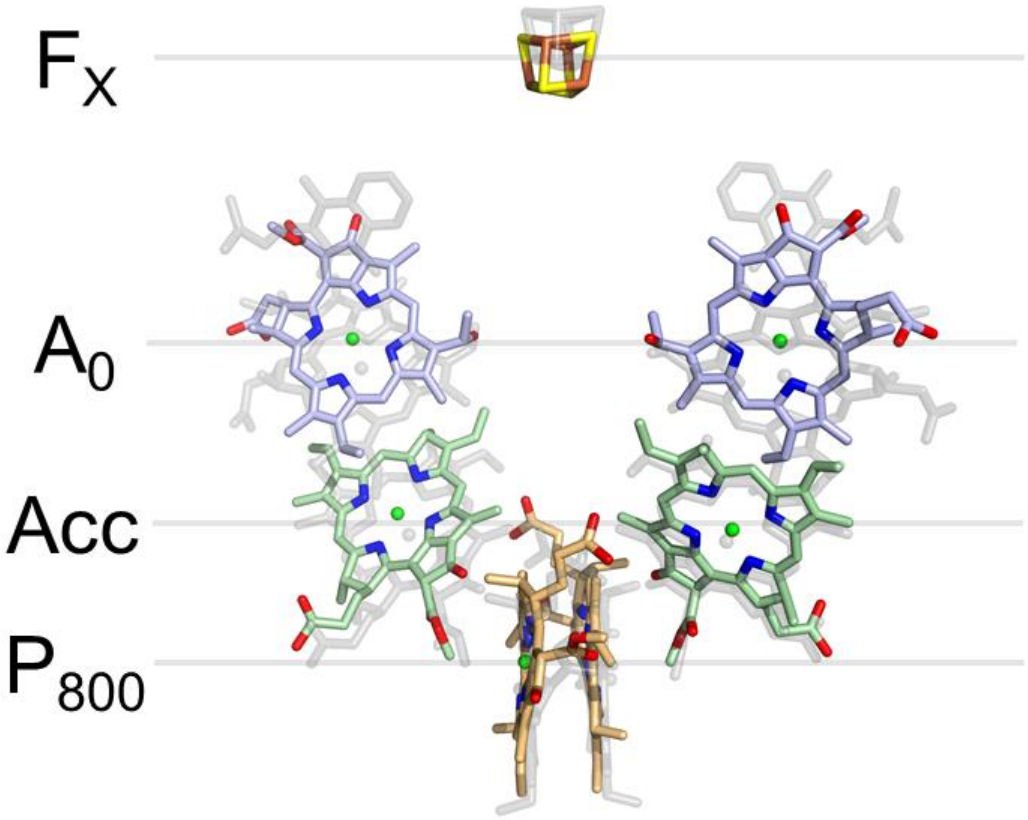


Figure A7 Superposition of the ET chains. PSI (transparent grey, PDB ID: 1JB0) and HbRC (colored) ET chains are superimposed. Transparent grey lines indicate the names of the (B)Chl in the corresponding position (left) and are in the same locations as those of Fig. 3.

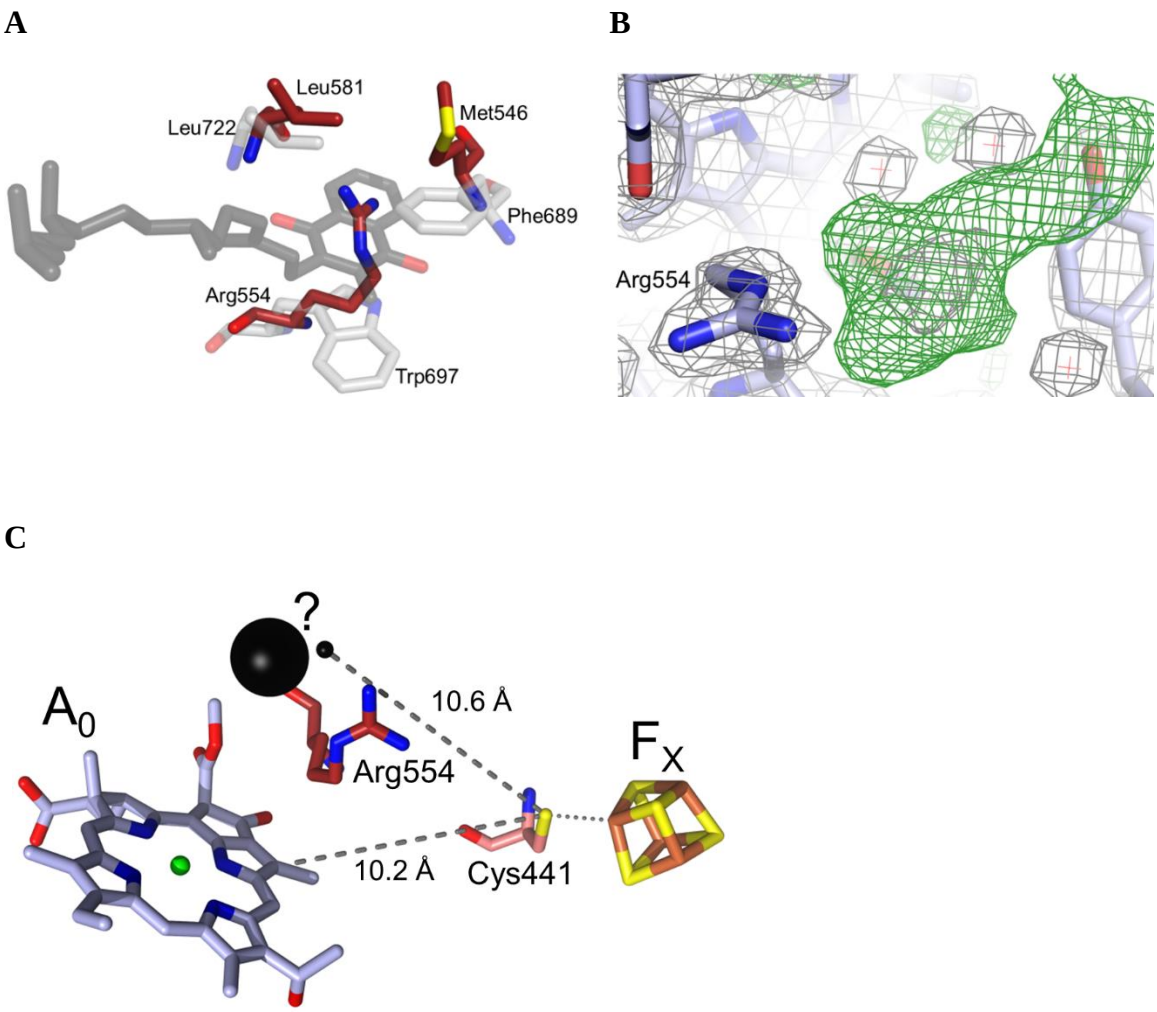


Figure A8 Binding of quinone to Type I RCs. (**A**) Phylloquinone (dark transparent grey carbons) in PSI and the three residues by which it is coordinated (light transparent grey carbons). The analogous residues from the HbRC (dark red carbons) are superimposed, including the Arg that occupies the space analogous to the quinone-binding site (PDB ID: 1JB0). (**B**) Unassigned electron density near A_0 within the HbRC electron density map. Green shows the F_O-F_C map at a contour level of 3σ and grey shows the $2F_O$-F_C map

(contour level of 2σ). (**C**) The center of mass of the unassigned density is shown as a large black sphere, while the closest edge of the density (map contour at 1.5σ) to F_X is shown as a small black sphere. The closest F_X-coordinating residue (Cys441) and the Arg located between the unassigned density and F_X (Arg554) are also shown. The distances shown on the figure are the closest distances to F_X (thiolate sulfur of Cys441) from either the unassigned density (closest point) or A_0 (carbon 12, part of the conjugated π system of the (B)Chl macrocycle).

Table A1 Expanded Crystallographic Table

Wavelength (Å)	1.009
Resolution range (Å)	29.08 - 2.20 (2.28 - 2.20)
Space group	C 1 2 1
Unit cell	a = 131.36 Å, b = 89.71 Å, c = 111.64 Å
	α = 90°, β = 108.15°, γ = 90°
Total reflections	118,884 (9681)
Unique reflections	61,415 (5534)
Multiplicity	1.9 (1.7)
Completeness (%)	98.2 (89.2)
Mean I/sigma(I)	16.6 (1.7)
Wilson *B* factor (Å^2)	41.36
R_{merge}	0.030 (0.48)
R_{meas}	0.042 (0.68)
R_{pim}	0.030 (0.48)
CC1/2 (%)	99.7 (71.8)

CC* (%)	99.9 (91.4)
Reflections used in refinement	
Reflections used for R-free	2936 (292)
R-work (%)	15.85 (25.68)
R-free (%)	19.08 (30.71)
CC(work) (%)	95.7 (87.9)
CC(free) (%)	93.2 (80.3)
Number of non-hydrogen atoms	6950
macromolecules	4898
ligands	1812
solvent	240
Protein residues	626
RMS(bonds) (Å)	0.012
RMS(angles) (°)	1.57
Ramachandran favored (%)	98

Ramachandran allowed (%)	2.1
Ramachandran outliers (%)	0.16
Rotamer outliers (%)	0.41
Clashscore	2.41
Average B-facto (Å2	54
macromolecules	53
ligands	58
solvent	64

Statistics for the highest-resolution bin are shown in parentheses.

Table A2 Distances between ET cofactors and calculated decay times of ET

Cofactor pair	Center to Center (Å)	Edge to Edge (Å)	τ_{ideal} of ET[1]
P-Acc	12.1	4.9	0.86 ps
P-A_0	19.4	12.1	18 ns
P-F_x	29.0	24.0	270 ms
Acc-A_0	9.1	3.3	96 fs
Acc-F_x	24.3	16.8	12 μs
A_0-F_x	17.9	10.2	1.3 ns

[1]Rates were calculated using the Moser-Dutton formula for the maximal rate of an ET reaction (i.e. if $\Delta G = -\lambda$): $\log_{10}(k_{ideal}) = 15 - 0.6R$, where R is the closest distance between the cofactors. Atoms used for edge-to-edge distances included any atom of a (B)Chl that was part of the conjugated π system, and any Fe or S atom of or ligated to F_X (including those of Cys thiolates) [14].

Table A3 Antenna Domain Resides Selected for TMH Superposition

Polypeptide	TMH1	TMH2	TMH3	TMH4	TMH5	TMH6
PshA	26-49	85-106	121-146	170-194	208-232	266-299
PsaA	65-96	155-182	193-229	294-313	352-376	387-419
PsaB	38-70	131-156	170-202	269-288	333-357	368-400
PsbB (CP47)	16-44	93-116	135-159	195-218	234-258	447-475
PsbC (CP43)	46-74	108-135	154-181	230-252	268-292	422-453

TMH residues selected for superposition of antenna domains compared to that of the HbRC structure (Fig. S5). Background denotes membership in Type I (green) or Type II (blue) RCs.

Table A4 ET Domain TMH Selection for Superposition

Polypeptide	TMH7/1	TMH8/2	TMH9/3	TMH10/4	TMH11/5
PshA	320-353	384-409	445-467	528-545	583-607
PsaA	437-468	533-559	591-621	669-691	724-754
PsaB	421-449	520-547	578-609	650-670	708-737
PsbA (D1)	32-54	110-137	143-165	196-220	271-295
PsbD (D2)	32-54	109-136	141-163	195-220	264-291
L	33-55	84-111	115-139	171-198	226-250
M	53-76	111-138	142-166	198-224	261-284

TMH residues selected for superposition of ET domains compared to that of the HbRC structure (Fig. S6). Background denotes membership in Type I (green) or Type II (blue) RCs.

APPENDIX B

SUPPLEMENTARY INFORMATION FOR CHAPTER 4

Figures

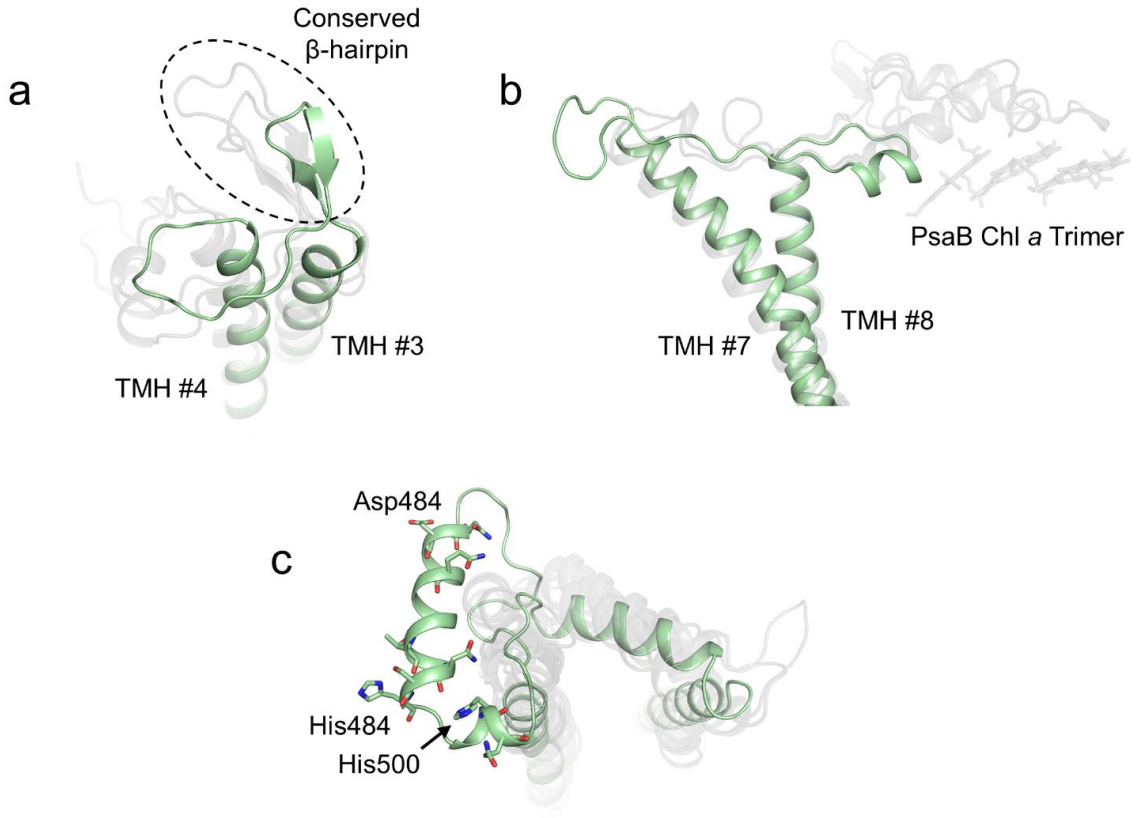


Figure B1 Surface comparisons of the HbRC with other RC structures. ET and antenna domains were superimposed based on their transmembrane helices. (**a**) A β-hairpin between TMH #3 and TMH #4 on the donor side of the HbRC occupies a similar position to those of CP43 and CP47 of PSII (transparent grey, structures extracted from PDB ID: 3WU2). (**b**) The loop between TMH7 and TMH8 on the donor side is smaller than in PSI (PDB ID: 1JB0) where this loop coordinated a dimer (PsaA, not shown) or trimer (PsaB, transparent grey) of chlorophylls known to be a site of red light absorbance (ref). (**c**) The

loop between TMH10 and TMH11 on the donor side that is known to contain a surface helix that is involved in donor binding (ref, labelled “P-helix”) is longer in the HbRC, with an additional two surface helices that contain charged (labelled) and uncharged polar residues. Analogous regions from structures of the L and M subunits of the PbRC (PDB ID: 1PRC), the D1 and D2 subunits of PSII (PDB ID: 3WU2), and PsaA and PsaB of PSI (PDB ID: 1JB0), are shown in transparent grey.

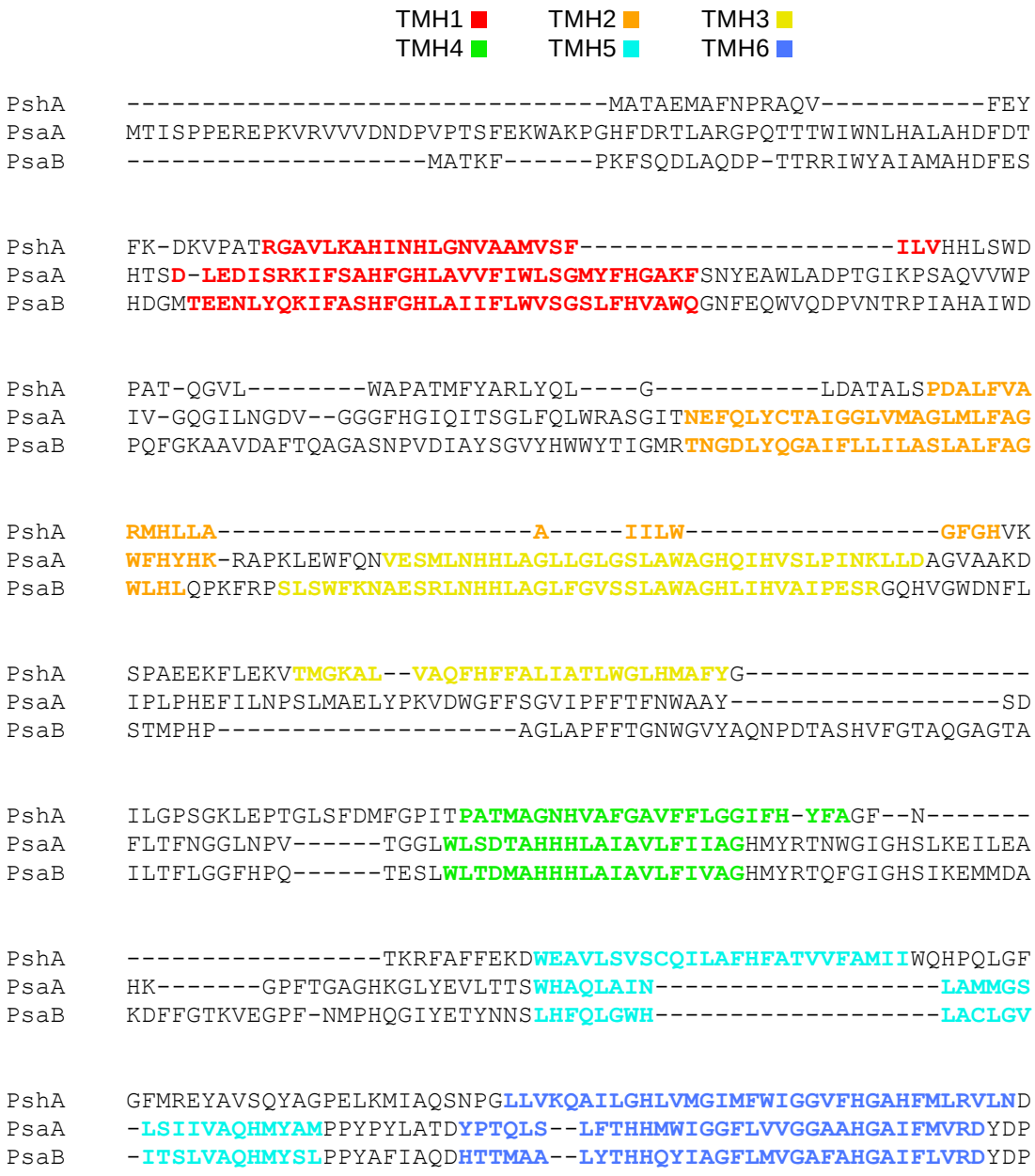


Figure B2 Non-structure-based alignment of PshA, PsaA, and PsaB antenna domains. Non-structure-based alignment (Clustal Omega) of antenna domains from PshA (*H. modesticaldum*, PDB code 5V8K), PsaA (*S. elongates*, PDB code 1JB0), and PsaB (*S. elongates*, PDB code 1JB0). Transmembrane helix residues identified from the crystal

structures are colored.

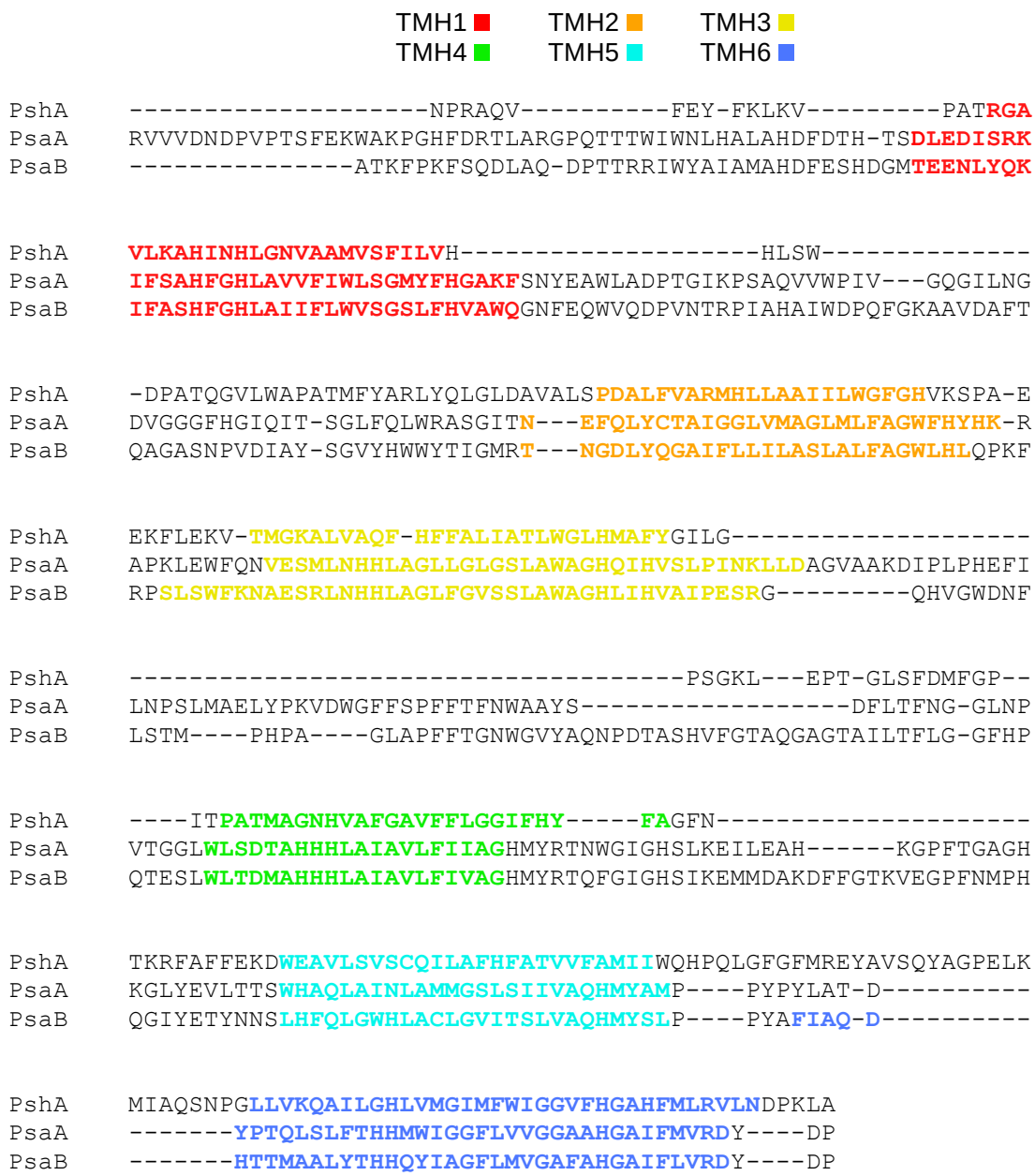


Figure B3 Structure-based alignment of PshA, PsaA, and PsaB antenna domains. Structure-based alignment (PROMALS3D) of antenna domains from PshA (*H. modesticaldum*, PDB code 5V8K), PsaA (*S. elongates*, PDB code 1JB0), and PsaB (*S.*

elongates, PDB code 1JB0). Transmembrane helix residues identified from the crystal structures are colored. Note that the alignment produced in TMH 2, 3, and 5 contains fewer gaps than that from the non-structure-based alignment in Fig. S3.

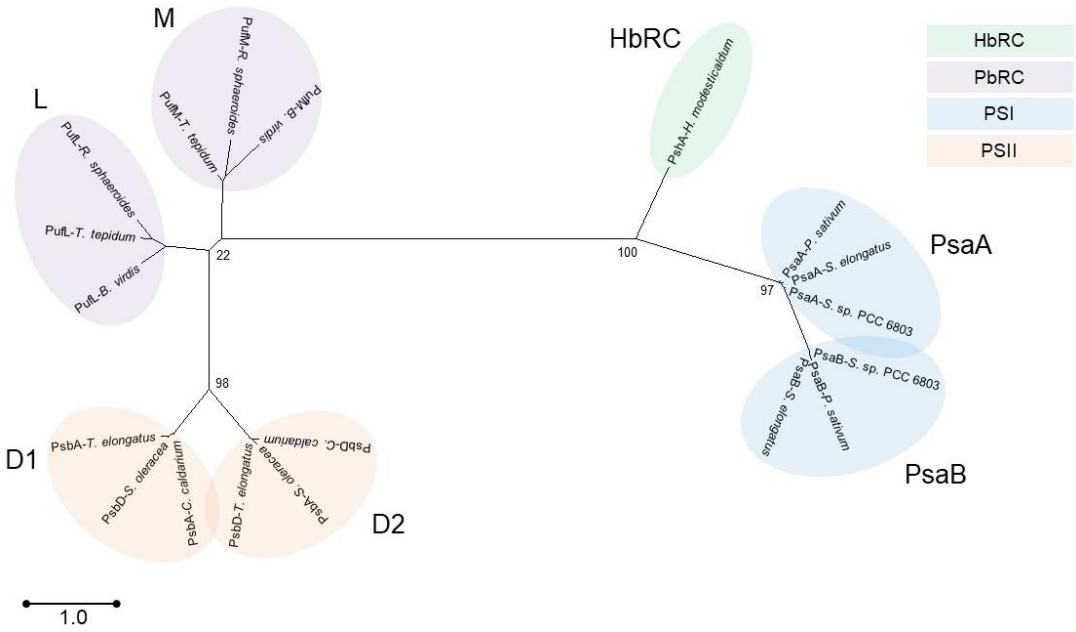
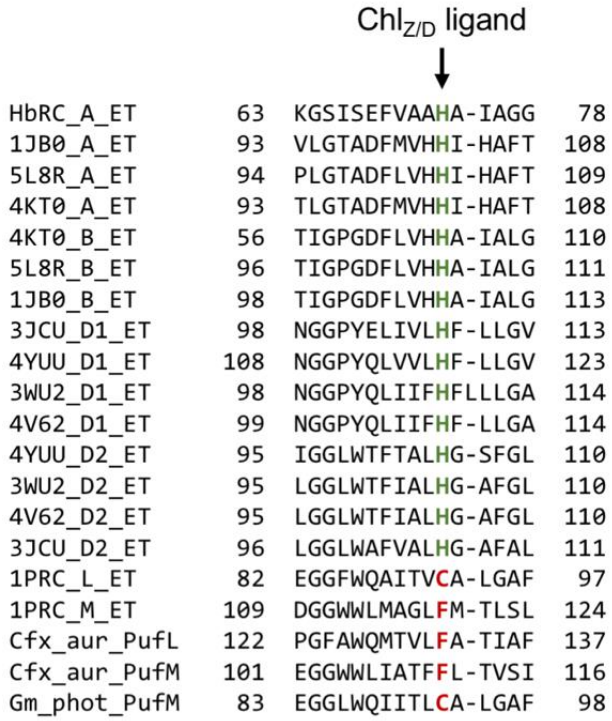


Figure B4 Structure-based phylogenetic tree of ET domains. This tree was created using structure-based (PROMALS3D) MSAs of the ET domain. The evolutionary history was inferred by using the Maximum Likelihood method based on the Le and Gascuel model (Le and Gascuel 2008). The tree with the highest log likelihood (-4629.76) is shown. Initial tree(s) for the heuristic search were obtained automatically by applying Neighbor-Join and BioNJ algorithms to a matrix of pairwise distances estimated using a JTT model, and then selecting the topology with superior log likelihood value. A discrete Gamma distribution was used to model evolutionary rate differences among sites (4 categories (+*G*, parameter = 6.7251)). The tree is drawn to scale, with branch lengths measured in

the number of substitutions per site (scale bar in bottom left). The analysis involved 19 amino acid sequences. All positions containing gaps and missing data were eliminated. There were a total of 182 positions in the final dataset. The percentage of trees in which in which the associated taxa clustered together in the bootstrap test (500 replicates) are shown next to the branches. The RC with which the antenna polypeptide is affiliated, and the node significance, are listed in the legend. Potential roots are discussed in the text using the letters “A” and “B”.

a



b

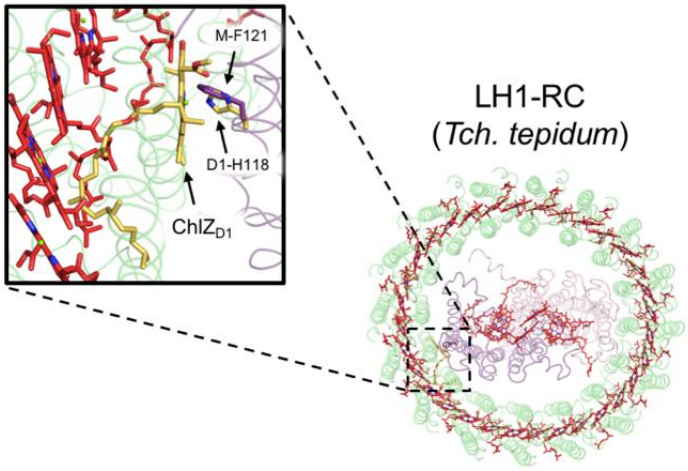


Figure B5 Comparison of the Chl$_{Z/D}$ binding sites amongst various RCs. (**a**) All RCs, except those in the PbRC lineage, have been shown to bind this ligand. Although no structures are available for the CfxRC or GmRC, sequence alignments show that these

RCs should not bind the ligand, like the rest of the PbRCs. (**b**) The structure of LH1-RC from *Tch. tepidum* (PDB code 3WMM), overlaid with the $ChlZ_{D1}$ ligand from PSII (PDB code 4V62), illustrating that $ChlZ_{D1}$ clashes with TMHs of LH1. For this overlay, D1/D2 was structurally aligned to L/M using the *super* function of Pymol. All elements of PSII were hidden from view, except for $ChlZ_{D1}$.

Table B1 Structure-based MSA of partial ET domains. The highlighted amino acid position lies on the P-Helix and faces the accessory chlorophyll.

Polypeptide	Species	Structure-based MSA
PshA	*H. modesticaldum* (5V8K)	505 LTN-------MGY-FSYIIQQTTAM 521
PsaA	*S. elongatus* (1JB0)	640 GNFAQSAITINGWLRDFLWAQASQV 661
PsaA	*S.* sp. PCC 6803 (4KT0)	633 GNFAQSAITINGWLRDFLWAQAANV 657
PsaA	*P. sativum* (5L8R)	640 GNFAQSSITINGWLRDFLWAQASQV 664
PsaB	*S. elongatus* (1JB0)	613 AQFNESSTYLMGWLRDYLWLNSSQL 637
PsaB	*S.* sp. PCC 6803 (4KT0)	604 AQFNENSTYLMGWFRDYLWANSAQL 628
PsaB	*P. sativum* (5L8R)	607 SQFNESSTYLMGWLRDYLWLNSSQL 631
PsbA	*T. elongatus* (4V62)	169 S-DG-MPLGISGTF-NFMIVFQAEH 190
PsbA	*C. caldarium* (4YUU)	169 S-DG-MPLGISGTF-NFMLVFQAEH 190
PsbA	*S. oleracea* (3JCU)	169 S-DG-MPLGISGTF-NFMIVFQAEH 190
PsbD	*T. elongatus* (4V62)	168 F-FA-PSFGVAAIF-RFLLFFQGFH 189
PsbD	*C. caldarium* (4YUU)	168 F-FA-PSFGVAAIF-RFLLFLQGFH 189
PsbD	*S. oleracea* (3JCU)	168 F-FA-PSFGVAAIF-RFILFFQGFH 189
PufL	*B. virdis* (1PRC)	143 G-HA-FPYGILSHL-DWVNNFGYQY 164
PufL	*R. sphaeroides* (1PCR)	143 G-YA-FPYGIWTHL-DWVSNTGYTY 164
PufL	*T. tepidum* (1EYS)	151 G-HG-FPYGILSHL-DWVSNVGYQF 172
PufM	*B. virdis* (1PRC)	170 S-EG-VPFGIWPHI-DWLTAFSIRY 191
PufM	*R. sphaeroides* (1PCR)	172 S-EA-VPYGIFSHL-DWTNNFSLVH 193

PufM	*T. tepidum* (1EYS)	171 A-KA-VPFGIFPHL-DWTAAFSIRY 192

Bacterial Type I, Oxygenic Type I, Oxygenic Type II, and Bacterial Type II polypeptides are colored in green, blue, tan, and purple, respectively.

Table B2 Antenna Domains Extracted from PDB Structures

Polypeptide Name	**Organism**	**PDB ID**	**Residue Range**
PshA	*Heliobacterium modesticaldum*	5V8K	9-318
PsaA	*Synechococcus elongatus*	1JB0	13-437
PsaA	*Synechocystis* sp. PCC 6803	4KT0	13-433
PsaA	*Pisum sativum*	5L8R	16-439
PsaB	*Synechococcus elongatus*	1JB0	1-419
PsaB	*Synechocystis* sp. PCC 6803	4KT0	3-413
PsaB	*Pisum sativum*	5L8R	2-415
PsbB	*Thermosynechococcus elongatus*	4V62	2-491
PsbB	*Cyanidium caldarium*	4YUU	2-484
PsbB	*Spinacia oleracea*	3JCU	2-488
PsbC	*Thermosynechococcus elongatus*	4V62	27-473
PsbC	*Cyanidium caldarium*	4YUU	25-473
PsbC	*Spinacia oleracea*	3JCU	24-472

Antenna domains associated with Type I and Type II RCs are shaded blue and tan, respectively.

Table B3 ET Domains Extracted from PDB Structures

Polypeptide Name	Organism	PDB ID	Residue Range
PshA	*Heliobacterium modesticaldum*	5V8K	
PsaA	*Synechococcus elongatus*	1JB0	438-755
PsaA	*Synechocystis* sp. PCC 6803	4KT0	434-751
PsaA	*Pisum sativum*	5L8R	440-758
PsaB	*Synechococcus elongatus*	1JB0	420-739
PsaB	*Synechocystis* sp. PCC 6803	4KT0	414-730
PsaB	*Pisum sativum*	5L8R	416-734
PsbB	*Thermosynechococcus elongatus*	4V62	10-344
PsbB	*Cyanidium caldarium*	4YUU	1-344
PsbB	*Spinacia oleracea*	3JCU	11-344
PsbC	*Thermosynechococcus elongatus*	4V62	13-352
PsbC	*Cyanidium caldarium*	4YUU	13-352
PsbC	*Spinacia oleracea*	3JCU	12-351
PufL	*Thermochromatium tepidum*	3WMM	2-281
PufL	*Blastochloris viridis*	1PRC	1-273
PufL	*Rhodobacter sphaeroides*	1PCR	1-281
PufM	*Thermochromatium tepidum*	3WMM	2-320
PufM	*Blastochloris viridis*	1PRC	1-323
PufM	*Rhodobacter sphaeroides*	1PCR	1-302

ET domains associated with Type I and Type II RCs are shaded blue and tan, respectively.

Table B4 Top 5 models of each MSA

Antenna Domain		
NSB (Clustal Omega)	**SB (PROMALS3D)**	**SB (PDBeFOLD)**
LG+G	cpREV+F	cpREV+G+F
LG+G+I	cpREV+G+I+F	cpREV+F
LG+I	LG+G+F	WAG+G+I+F
LG	LG+G+I+F	WAG+F
WAG+F	WAG+G+F	WAG+I+F

ET Domain		
NSB (Clustal Omega)	**SB (PROMALS3D)**	**SB (PDBeFOLD)**
LG+G+F	LG+G+F	WAG+G+F
LG+F	LG+G+I+F	WAG+G+I+F
rtREV+G+F	LG+F	WAG+I+F
LG+G+I+F	rtREV+G+F	WAG+F
LG+I+F	LG+I+F	LG+F

LG = Le and Gascuel Method, Green

WAG = Whelan and Goldman Method, Tan

cpREV = Plastid-Encoded Protein Method, Blue

rtREV = Reverse transcriptase Protein Method, Red

G = Gamma-distributed rate with 4 discrete gamma categories

I = Has invariant sites

F = With frequencies

Table B5 Matrix of substitution models and MSA trimming strategies for the 45 phylogenetic trees created for each domain.

Antenna Domain

NSB with cpREV+G+F	**PROMALS3D with cpREV+G+F**	**PDBeFOLD with cpREV+G+F**
Complete Deletion	Complete Deletion	Complete Deletion
75% Site Coverage Cutoff	75% Site Coverage Cutoff	75% Site Coverage Cutoff
50% Site Coverage Cutoff	50% Site Coverage Cutoff	50% Site Coverage Cutoff
25% Site Coverage Cutoff	25% Site Coverage Cutoff	25% Site Coverage Cutoff
All sites	All sites	All sites
NSB with WAG+G+I+F	**PROMALS3D with WAG+G+I+F**	**PDBeFOLD with WAG+G+I+F**
Complete Deletion	Complete Deletion	Complete Deletion
75% Site Coverage Cutoff	75% Site Coverage Cutoff	75% Site Coverage Cutoff
50% Site Coverage Cutoff	50% Site Coverage Cutoff	50% Site Coverage Cutoff
25% Site Coverage Cutoff	25% Site Coverage Cutoff	25% Site Coverage Cutoff
All sites	All sites	All sites
NSB with LG+G+F	**PROMALS3D with LG+G+F**	**PDBeFOLD with LG+G+F**
Complete Deletion	Complete Deletion	Complete Deletion

75% Site Coverage Cutoff	75% Site Coverage Cutoff	75% Site Coverage Cutoff
50% Site Coverage Cutoff	50% Site Coverage Cutoff	50% Site Coverage Cutoff
25% Site Coverage Cutoff	25% Site Coverage Cutoff	25% Site Coverage Cutoff
All sites	All sites	All sites

ET Domain

NSB with LG+G+F	**PROMALS3D with LG+G+F**	**PDBeFOLD with LG+G+F**
Complete Deletion	Complete Deletion	Complete Deletion
75% Site Coverage Cutoff	75% Site Coverage Cutoff	75% Site Coverage Cutoff
50% Site Coverage Cutoff	50% Site Coverage Cutoff	50% Site Coverage Cutoff
25% Site Coverage Cutoff	25% Site Coverage Cutoff	25% Site Coverage Cutoff
All sites	All sites	All sites
NSB with rtREV+G+F	**PROMALS3D with rtREV+G+F**	**PDBeFOLD with rtREV+G+F**
Complete Deletion	Complete Deletion	Complete Deletion
75% Site Coverage Cutoff	75% Site Coverage Cutoff	75% Site Coverage Cutoff
50% Site Coverage Cutoff	50% Site Coverage Cutoff	50% Site Coverage Cutoff
25% Site Coverage Cutoff	25% Site Coverage Cutoff	25% Site Coverage Cutoff
All sites	All sites	All sites

NSB with WAG+G+F	**PROMALS3D with WAG+G+F**	**PDBeFOLD with WAG+G+F**
Complete Deletion	Complete Deletion	Complete Deletion
75% Site Coverage Cutoff	75% Site Coverage Cutoff	75% Site Coverage Cutoff
50% Site Coverage Cutoff	50% Site Coverage Cutoff	50% Site Coverage Cutoff
25% Site Coverage Cutoff	25% Site Coverage Cutoff	25% Site Coverage Cutoff
All sites	All sites	All sites